carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Particularly, N-terminal deletions of the polypeptide of the present invention can be described by the general formula m-p, where p is the total number of amino acids in the polypeptide and m is an integer from 2 to (p-1), and where both of these integers (m & p) correspond to the position of the amino acid residue identified in SEQ ID NO:Y.

Moreover, C-terminal deletions of the polypeptide of the present invention can also be described by the general formula 1-n, where n is an integer from 2 to (p-1), and again where these integers (n & p) correspond to the position of the amino acid residue identified in SEQ ID NO:Y.

The invention also provides polypeptides having one or more amino acids deleted from both the amino and the carboxyl termini, which may be described generally as having residues m-n of SEQ ID NO:Y, where m and n are integers as described above.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an

epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the

polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D.

Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance

genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein

after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

5

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

10 The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

15 Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing 20 the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping 25 strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This 30 technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to 35 mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are

more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library) .) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model

systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of

unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic

resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activity

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can

decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitis, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic

shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstron's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases

may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiollitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox , hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), Pasteurellacea Infections (e.g., Actinobacillus, Heamophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS

related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteocarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal

or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stoke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

Chemotaxis

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotaxic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotaxic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotaxic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors),or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity , and (b) determining if a biological activity of the polypeptide has been altered.

Other Activities

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, caricadic rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method

comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95%

identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide

comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

Vector Used to Construct Library	Corresponding Deposited Plasmid
Lambda Zap	pBluescript (pBS)
Uni-Zap XR	pBluescript (pBS)
Zap Express	pBK
lafmid BA	plafmid BA
pSport1	pSport1
pCMVSport 2.0	pCMVSport 2.0
pCMVSport 3.0	pCMVSport 3.0
pCR®2.1	pCR®2.1

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1

Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.

Vectors pSport1, pCMVSport 2.0 and pCMVSport 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 Blue. Vector pCR®2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with ^{32}P-γ-ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).)

The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 µl of reaction mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is 1.5-5 mM $MgCl_2$, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

Example 3: Tissue Distribution of Polypeptide

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P^{32} using the rediprime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

Example 4: Chromosomal Mapping of the Polynucleotides

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This primer set is then used in a polymerase chain reaction under the following set of conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

Example 5: Bacterial Expression of a Polypeptide

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 ($O.D.^{600}$) of between 0.4 and 0.6. IPTG

(Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM immidazole. Immidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4° C or frozen at -80° C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA

insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

Example 6: Purification of a Polypeptide from an Inclusion Body

The following alternative method can be used to purify a polypeptide expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfuidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area

(e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 μg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., Virology 170:31-39 (1989).

Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five μg of a plasmid containing the polynucleotide is co-transfected with 1.0 μg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One μg of BaculoGold™ virus DNA and 5 μg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 μl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 μl Lipofectin plus 90 μl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm

tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 µCi of ^{35}S-methionine and 5 µCi ^{35}S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

Example 8: Expression of a Polypeptide in Mammalian Cells

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates

the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0, and pCMVSport 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No.209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the

polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five µg of the expression plasmid pC6 is cotransfected with 0.5 µg of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of metothrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 µM, 2 µM, 5 µM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 µM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Example 9: Protein Fusions

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

Human IgG Fc region:

GGGATCCGGAGCCCAAATCTTCTGACAAAACTCACACATGCCCACCGTGCC
CAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCCAAAACC
CAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTCACATGCGTGGTGGT
GGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACG
GCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAAC

AGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTG
AATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAACCCCC
ATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGT
GTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCT
GACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGTGGAGTGGGA
GAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGG
ACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCA
GGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGC
ACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAGTGC
GACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

Example 10: Production of an Antibody from a Polypeptide

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as

described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

Example 11: Production Of Secreted Protein For High-Throughput Screening Assays

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a

working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2 x 10^5 cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl2 (anhyd); 0.00130 mg/L $CuSO_4$-$5H_2O$; 0.050 mg/L of $Fe(NO_3)_3$-$9H_2O$; 0.417 mg/L of $FeSO_4$-$7H_2O$; 311.80 mg/L of Kcl; 28.64 mg/L of $MgCl_2$; 48.84 mg/L of $MgSO_4$; 6995.50 mg/L of NaCl; 2400.0 mg/L of $NaHCO_3$; 62.50 mg/L of NaH_2PO_4-H_20; 71.02 mg/L of Na_2HPO4; .4320 mg/L of $ZnSO_4$-$7H_2O$; .002 mg/L of Arachidonic Acid ; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitric Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of

Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-Glucose; 130.85 mg/ml of L- Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H_20; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL-H_20; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H_20; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalainine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tryrosine-2Na-$2H_20$; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319 mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B_{12}; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine; 0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

Example 12: Construction of GAS Reporter Construct

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proxial region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

Ligand	JAKs tyk2	Jak1	Jak2	Jak3	STATS	GAS(elements) or ISRE
IFN family						
IFN-a/B	+	+	-	-	1,2,3	ISRE
IFN-g		+	+	-	1	GAS (IRF1>Lys6>IFP)
Il-10	+	?	?	-	1,3	
gp130 family						
IL-6 (Pleiotrohic)	+	+	+	?	1,3	GAS (IRF1>Lys6>IFP)
Il-11(Pleiotrohic)	?	+	?	?	1,3	
OnM(Pleiotrohic)	?	+	+	?	1,3	
LIF(Pleiotrohic)	?	+	+	?	1,3	
CNTF(Pleiotrohic)	-/+	+	+	?	1,3	
G-CSF(Pleiotrohic)	?	+	?	?	1,3	
IL-12(Pleiotrohic)	+	-	+	+	1,3	
g-C family						
IL-2 (lymphocytes)	-	+	-	+	1,3,5	GAS
IL-4 (lymph/myeloid)	-	+	-	+	6	GAS (IRF1 = IFP >>Ly6)(IgH)
IL-7 (lymphocytes)	-	+	-	+	5	GAS
IL-9 (lymphocytes)	-	+	-	+	5	GAS
IL-13 (lymphocyte)	-	+	?	?	6	GAS
IL-15	?	+	?	+	5	GAS
gp140 family						
IL-3 (myeloid)	-	-	+	-	5	GAS (IRF1>IFP>>Ly6)
IL-5 (myeloid)	-	-	+	-	5	GAS
GM-CSF (myeloid)	-	-	+	-	5	GAS
Growth hormone family						
GH	?	-	+	-	5	
PRL	?	+/-	+	-	1,3,5	
EPO	?	-	+	-	5	GAS(B-CAS>IRF1=IFP>>Ly6)
Receptor Tyrosine Kinases						
EGF	?	+	+	-	1,3	GAS (IRF1)
PDGF	?	+	+	-	1,3	
CSF-1	?	+	+	-	1,3	GAS (not IRF1)

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:
5':GCGCCTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCG
AAATGATTTCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:
5':CTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCGAAATG
ATTTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCC
CTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGC
CCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGC
CTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTT
TGCAAAAAGCTT:3' (SEQ ID NO:5)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, Il-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

Example 13: High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml genticin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1%Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies)

with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells (10^7 per transfection), and resuspend in OPTI-MEM to a final concentration of 10^7 cells/ml. Then add 1ml of 1 x 10^7 cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Genticin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100, 000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

Example 14: High-Throughput Screening Assay Identifying Myeloid Activity

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest $2x10e^7$ U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM $Na_2HPO_4.7H_2O$, 1 mM $MgCl_2$, and 675 uM $CaCl_2$. Incubate at 37°C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

These cells are tested by harvesting $1x10^8$ cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of $5x10^5$ cells/ml. Plate 200 ul cells per well in the 96-well plate (or $1x10^5$ cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat phenochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)

5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine

growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as $5x10^5$ cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to $1x10^5$ cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

Example 16: High-Throughput Screening Assay for T-cell Activity

NF-κB (Nuclear Factor κB) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF-κB regulates the expression of genes involved in immune cell activation, control of apoptosis (NF-κB appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- κB is retained in the cytoplasm with I-κB (Inhibitor κB). However, upon stimulation, I- κB is phosphorylated and degraded, causing NF- κB to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- κB include IL-2, IL-6, GM-CSF, ICAM-1 and class 1 MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF-κB promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF-kB would be useful in treating

diseases. For example, inhibitors of NF-κB could be used to treat those diseases related to the acute or chronic activation of NF-kB, such as rheumatoid arthritis.

To construct a vector containing the NF-κB promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF-κB binding site (GGGGACTTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:
5':GCGGCCTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGAC
TTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:
5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene) Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

5':CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGACTTTCC
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCA
TCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACT
AATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTC
CAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTT:
3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2-promoter plasmid (Clontech) with this NF-κB/SV40 fragment using XhoI and HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF-κB/SV40/SEAP cassette is removed from the above NF-κB/SEAP vector using restriction enzymes SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly, the NF-κB/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

Once NF-κB/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1,1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 µl of 2.5x dilution buffer into Optiplates containing 35 µl of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 µl Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 µl Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

Reaction Buffer Formulation:

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6

23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25
28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-3, used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO_2 incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-3 is made in 10% pluronic acid DMSO. To load the cells with fluo-3, 50 ul of 12 ug/ml fluo-3 is added to each well. The plate is incubated at 37°C in a CO_2 incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10^6 cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-3 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10^6 cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-3. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca^{++} concentration.

Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford,MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford,MA) are used to cover the Loprodyne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na3VO4, 2 mM Na4P2O7 and a cocktail of protease inhibitors (# 1836170) obtained from Boeheringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a

biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg_{2+} (5mM ATP/50mM $MgCl_2$), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM $MgCl_2$, 5 mM $MnCl_2$, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30°C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mm EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavadin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phospotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above.

Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity

As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other

phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies).

The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Manheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10.

The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 23: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 µg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 µg/kg/hour to about 50 µg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracistemally, intravaginally,

intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), bucally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or mirocapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2- hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 24: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

Example 25: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

Example 26: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

Example 27: Method of Treatment Using Gene Therapy - In Vivo

Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide of the present invention. A polynucleotide of the present invention may be operatively linked to a promoter or any other genetic elements necessary for the expression of the encoded polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO90/11092, WO98/11779; U.S. Patent NO. 5693622, 5705151, 5580859; Tabata H. et al. (1997) Cardiovasc. Res. 35(3):470-479, Chao J et al. (1997) Pharmacol. Res. 35(6):517-522, Wolff J.A. (1997) Neuromuscul. Disord. 7(5):314-318, Schwartz B. et al. (1996) Gene Ther. 3(5):405-411, Tsurumi Y. et al. (1996) Circulation 94(12):3281-3290 (incorporated herein by reference).

The polynucleotide constructs of the present invention may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). These polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides may also be delivered in liposome formulations (such as those taught in Felgner P.L. et al. (1995) Ann. NY Acad. Sci. 772:126-139 and Abdallah B. et al. (1995) Biol. Cell 85(1):1-7) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs of the present invention used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct of the present invention can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 g/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle *in vivo* is determined as follows. Suitable template DNA for production of mRNA coding for the polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with

liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice. The results of the above experimentation in mice can be use to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA of the present invention.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference.

Sequence Listing

(1) GENERAL INFORMATION:

(i) APPLICANT: Human Genome Sciences, Inc., et al.

(ii) TITLE OF INVENTION: 207 Human Secreted Proteins

(iii) NUMBER OF SEQUENCES: 800

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Human Genome Sciences, Inc.

(B) STREET: 9410 Key West Avenue

(C) CITY: Rockville

(D) STATE: Maryland

(E) COUNTRY: USA

(F) ZIP: 20850

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

(B) COMPUTER: HP Vectra 486/33

(C) OPERATING SYSTEM: MSDOS version 6.2

(D) SOFTWARE: ASCII Text

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE:

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Kenley K. Hoover

(B) REGISTRATION NUMBER: 40,302

(C) REFERENCE/DOCKET NUMBER: PZ007PCT

(vi) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (301) 309-8504

(B) TELEFAX: (301) 309-8439

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 733 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGGATCCGGA GCCCAAATCT TCTGACAAAA CTCACACATG CCCACCGTGC CCAGCACCTG 60

AATTCGAGGG TGCACCGTCA GTCTTCCTCT TCCCCCCAAA ACCCAAGGAC ACCCTCATGA 120

TCTCCCGGAC TCCTGAGGTC ACATGCGTGG TGGTGGACGT AAGCCACGAA GACCCTGAGG 180

TCAAGTTCAA CTGGTACGTG GACGGCGTGG AGGTGCATAA TGCCAAGACA AAGCCGCGGG 240

AGGAGCAGTA CAACAGCACG TACCGTGTGG TCAGCGTCCT CACCGTCCTG CACCAGGACT 300

GGCTGAATGG CAAGGAGTAC AAGTGCAAGG TCTCCAACAA AGCCCTCCCA ACCCCCATCG 360

AGAAAACCAT CTCCAAAGCC AAAGGGCAGC CCCGAGAACC ACAGGTGTAC ACCCTGCCCC 420

CATCCCGGGA TGAGCTGACC AAGAACCAGG TCAGCCTGAC CTGCCTGGTC AAAGGCTTCT 480

ATCCAAGCGA CATCGCCGTG GAGTGGGAGA GCAATGGGCA GCCGGAGAAC AACTACAAGA 540

CCACGCCTCC CGTGCTGGAC TCCGACGGCT CCTTCTTCCT CTACAGCAAG CTCACCGTGG 600

ACAAGAGCAG GTGGCAGCAG GGGAACGTCT TCTCATGCTC CGTGATGCAT GAGGCTCTGC 660

ACAACCACTA CACGCAGAAG AGCCTCTCCC TGTCTCCGGG TAAATGAGTG CGACGGCCGC 720

GACTCTAGAG GAT 733

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Trp Ser Xaa Trp Ser
1 5

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 86 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GCGCCTCGAG ATTTCCCCGA AATCTAGATT TCCCCGAAAT GATTTCCCCG AAATGATTTC 60

CCCGAAATAT CTGCCATCTC AATTAG 86

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GCGGCAAGCT TTTTGCAAAG CCTAGGC 27

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 271 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CTCGAGATTT CCCCGAAATC TAGATTTCCC CGAAATGATT TCCCCGAAAT GATTTCCCCG 60

AAATATCTGC CATCTCAATT AGTCAGCAAC CATAGTCCCG CCCCTAACTC CGCCCATCCC 120

GCCCCTAACT CCGCCCAGTT CCGCCCATTC TCCGCCCCAT GGCTGACTAA TTTTTTTTAT 180

TTATGCAGAG GCCGAGGCCG CCTCGGCCTC TGAGCTATTC CAGAAGTAGT GAGGAGGCTT 240

TTTTGGAGGC CTAGGCTTTT GCAAAAAGCT T 271

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GCGCTCGAGG GATGACAGCG ATAGAACCCC GG 32

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GCGAAGCTTC GCGACTCCCC GGATCCGCCT C 31

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGGGACTTTC CC 12

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 73 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GCGGCCTCGA GGGGACTTTC CCGGGGACTT TCCGGGGACT TTCCGGGACT TTCCATCCTG 60

CCATCTCAAT TAG 73

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 256 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CTCGAGGGGA CTTTCCCGGG GACTTTCCGG GGACTTTCCG GGACTTTCCA TCTGCCATCT 60

CAATTAGTCA GCAACCATAG TCCCGCCCCT AACTCCGCCC ATCCCGCCCC TAACTCCGCC 120

CAGTTCCGCC CATTCTCCGC CCCATGGCTG ACTAATTTTT TTTATTTATG CAGAGGCCGA 180

GGCCGCCTCG GCCTCTGAGC TATTCCAGAA GTAGTGAGGA GGCTTTTTTG GAGGCCTAGG 240

CTTTTGCAAA AAGCTT 256

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2526 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GACAGGCTAT CCGAGAATCT GAGAGCTGGG CCCGGCAATT CCTCCAGYTA CCCTTGTGAC 60

CTAAGTCCAG TCACACATTT CCCAAAGTTT CTCTTTGTCA TAACCCTGGT CTGGCTGGTT 120

TTGRGGRCTT GAGAATGGGT CAGGGACTCC AGGCCAAGTC CAACAGAGAC CCCAAACCCA 180

CCACACACCA GCAGCCACAA CCTCACCACC AACAAAGAGG ACTTTTGTGG GGCCACAAGT 240

AAGAGGTCAT TTCTGGAATG GACTCAGACC TTTAAACAGG AGAGTTGAGC ACTTCCAGKS 300

AGTTTTTAAG CAAGGCATGG GGAACAGGGA ATAGAACCTT TCAAAGAGGT TGCCCAGAGA 360

AAAGCTGGGC CTCTTGCATT CGGCTTCCTT GGAGCAGCCT CTTCTGGCAG AAAGCCATCA 420

GGTGCTCAAT CATCTTCTCC TGGCCAAGGC TCTGACCATG CTTAGTACTG GAATAGAGGT 480

GGCCAGGCCC CCAGCGACTC TTCTTGGCCT GATGTTTGTC CTCACAGGCA TGCCACGTGG 540

CCTGAGATGA TTCAGAACAA ATCATGCTAA CTTTGAATCC ATCCAGCCAC TTGCAAATGA 600

TAATCAGAAG TCAGCTTGTT CACTGTTAGA AAGAAACTAA CAAAAGAGAA CCCAGAGCAA 660

TCTAGAATCT TTGAGTGCTT GGCTTTCCAA GGATACTGCG GAGACTCTGG CCAAGCTGAT 720

GAMCTTCTGA ARTGTCACTG GCACCATATG CAACAAGAAC CACCATTCAC TGAGTAGCTA 780

ATGGGTTTGG GGCCTGGGAC ATTCCATCTG AGGTCCTTCC TGAACATGTC ACTCCACAGC 840

AGAGGACCGG TTGCAGCTTA CCCAGAACCA CTCCTCCAGG AGAGCTGGAT GTTTTGCGTG 900

CAACACCTTG AGCACTGACT GCTATTGTTC AAAAAAAGCC TTTGCTGCAT TCGGAGGACT 960

GCCCCGTGCC CTGAGGTGAC TTCCTAACTA TGTGGTTTCA TTAGCGAATT TATTTTTTGT 1020

GCTGGGTGGA CATTTGTATT TTGTTAGGTT GCTGTTTAAG CTCAAGTTTG CTGTGCTCTC 1080

TGCAGCTACA AAACATCTTG GCATATTTAA GAKTGGCTTT TATAAATAGC TTTATTCTGA 1140

TATTAATCAG ATTCCCAACT TTACTGAGAA TTAAGGACTG GGGTACTTTA AAGAAATGCA 1200

AATAGCAATT GAAGAACCAC TGCTGCAGGT GGTAGCCCTG GCTAGACTGA ATTACACTAG 1260

AAATCAGCCA GAAGGAAGCG TCCTTGGGAT CCCAGATCAC TCTTTTTTTT TTTTTTTTTA 1320

AAAGGGGCAG CCCCTTGATG GCTCATCTCT CTGAATAACA GTTACGTCTT CATATCGATA 1380

CCAGATGCCT TCTTCATCAT GCCACTGAAG CCACTCACCA CCTTCAAGAA CATGCCAACC 1440

TCTGTCAGAT TCACTTACCC ACAAACAAGG AGGCACGTTT GGCACAAAGT GTTGTCCTCC 1500

AGGTCCAAGT GGACTCTACA GAGTGCTTGA CCTCAACACA CTGGATTCCA GGTGGACTGG 1560

ACCAAGAGCA GGCAAAGACA CGGGAACTGA AAAACTCCAC AGGGTTTGGA GAATAGAAAT 1620

GAAAAGCCAC GTCATATAAC TCAAGAATAA ATGGTGTTTT GGAAATTTTA AAATTATCAT 1680

CGAAGGTGGT GAAACTATTT CAGGCCCAAA TGAAAGGAAA TCGCCAGTTG GGGATGAAAT 1740

CACAGAGCCT GTGTTTTATG ATATGGTTGG ATGTCCACTG ATGAAATTTT AAAGGAGTTT 1800

CATTTTTAAA AGTGCGCATG ATTCTACATA TGAGAATTCT TTAGGCCAAG AAACTGTCCT 1860

TGGCTCAGAG GTGTTGGGAA TTAAAGCAGA GAGAAGCCAT TCGTGATGCT TAGAACCAAG 1920

GATGGTCATG TACACAAAGA CCATCGAGAC GGCCATTCTT GTTTACAAAA CACTTACCAA 1980

GAAAGCACTT TGTAGGGGAA CTTTAGTAAG TTCTTCTCAT TTCATTATGT TTCTTCCAAG 2040

GAAACAGGAG AGACTGAATT AATAATTCTC TCTTTCCTCT TAAGCACTTT TAAAATAATA 2100

AAGTACATCT TGAAATTTGG GGGGGCATCT CTGATTTAAA AAAAGAAAAA GGCTGCTTGA 2160

TGTATGTTAT GCAGAGACAC TCTGCCTCTG GTGGCTGCAG AGCAATACCC AAGCCTCATT 2220

TGGAAGGCTC AACATTTGGA ATTGCACTTT AATTGATTAA TCCTCAATTC ATGTGGCCTT 2280

ACGGGATGGT GGGTCTGGGA CCCCAATTCA TTCTTATCTG CCAAAGAATT ATCTAGAAGC 2340

ACATCAAATA CCAGCACCCC ACCTGCACAA TGGGGGTGGA AAACTTTTGT ATCCCTAAGC 2400

ATATTATTTT ATAGTGTCTG CCATGCCATG TGGAAATACT TTATTTTTAA CCTCAGGATT 2460

TAAATAAAGT AAACACTATG ACATTTAAAA AAAAAAAAAA AAAACTCGAG GGGGGCCCGG 2520

TACCCA 2526

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1131 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

CACTGCACCA GCTTTGTTAT CTGTAAAATG ATGATAATAC CAACACCTTC TTCTTGGGGT 60

ACTGAAGATG AGAGAACATG ATATGTGTAA AGTGCCTTCC ACAATACCCA GAACATAGCA 120

AACATGTAAT GAATGTAGTA ATAGTAATTA TTTTATTTTC TTTTGATTCA GTTGGGACTA 180

TGTTCAGCTG TAACAGAATA CCCAAAATAA CTGTTTTAAA CAAATTAAAG TTTWGTTGTG 240

AAGTTTTGTT ACGAATTCAG ACAATCCAGG GCTTTTATAG ATGCACCAGG ATCAGCAGGT 300

ACAAAGGCAT CTTTCCTGAT TTCTGCCAGT CTCAATGCAT GGGTTGCAAT CCAGARTCCA 360

RGATGGCAGT TCCAGCCCTG GTTACGCCCA TATTAGCACA CAGAAAGAAA GAGAAAGGGA 420

TGTGCCTCTT CACTTTAATC ATAGCTCCCA CTAGATGCAC CCACTACTTC TGCTGATACT 480

CCATTAGCTA ATGCTTGCTT ACATGGTCAC ACTTAGTTTC CAGAGAGACA TGTCTGGACA 540

GTCATGTGCT CAATTAATAT CCAAGTGTCC AATTACTGAG AAAAAAAGAA ACTAGCACCT 600

TTGCTTGGTT GCATTCCTCT TAGCATAAGC CACATTCTTT TTATGAAGTT GTCCTCAGTT 660

ACTTGGATGC CTCAGTTGTC CTTTCAWTTA GAAAWGCYCC TKGGACAYCC TGAAWCTGAC 720

TTCTTTTGTC ATCAGCACCA TCACTACCAC TGCCYTCTTC AAAGCCACCA CGTTCTGTCC 780

CCAGGATGGT TGCAACAACC ACCATAGGGA CTTTTTGCCT TCTACTTCCA CACAATAGNC 840

CAGAGTAAGC TTTTGAAAAT GTAGGTCAGA TCATGTCTCT CTCTTCCTCT TCAAAACCCT 900

CCCGATGGCT TTTCATATTA CTCAAAAGAA AACCTAAAAC TTTGCTGTGA GATCTATGTG 960

ACCCGGCTTA TTCTTCCTCT TACTTTATCT CTGTATTGCT CTTCCTCACT CTACTCCAGC 1020

CATCCCACCT CCTTGCTGCT TGTCCTATAC TCCTAAAAGA AGTTCAGTCT TCCCTTATGA 1080

TATTTGCACT TAAAATAGAA AAAAAAAAAA AAAAAAAACT CGAGGGGGGC C 1131

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 941 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GGCACGAGTA GCATTTCATT TAATCTGCAG GTATATTCTC CCAACAGTTT ATTGTCATGT 60

GATGTCCTCA GCCAAGATTG TRAGGCAGAG AGGAGCTGTC CCAACCTACT ATACCACCGA 120

GGCTGGAGAG ATCATATTTT TGGTATTAAA CTGGAGTCTC TCCATCCTTC ACATTGTTGA 180

TGTCCTCTGT AGCAAACCGG AAAAGTCAGT GACAGAAGAT GCCGCTAGCG GTTTGAGCCA 240

GAGAATGACA GCTCTGGTTT GGAGAAAAGG GCCGGATGGT GGCTCTAGAA AGCCCATCCT 300

TCTGCTCTTC TTTTTTCTCC CCCTTATATT GTGCTTTCAT TCATTCATTC ATTCATCAAA 360

CATTTGTTGA GCACCTATTA TGTGTCAAGC TCTGTGCTAG CCTCTGGAAA ACCTGCCCTC 420

ATGTAGCTCA CTGTGGAGTA GGAGAAACAA TGACTACACT ATGATAAGCA CGGGTTGTCA 480

GGGTCTCACA GAGCAGTGGC CCCTCATCCA GACCGATGAG GTCAAAGAAG GCATCCAGGC 540

GAGGATGGTG TCAGAGCTAA CTGAAGAATG AGAGGGAGCT GCACCASCAG GGGTTGGAAC 600

TGAAGGTGGC AGTGCCTGGA GTCTTGATTC CAGCAGAGGG AGAGCAGTCT GTGAAAAGGC 660

ACCAAGGGTG GGAGAGGGCA GAGCACATGG AGGAACTTCA GGTAGTTCTG GATGGCSCTG 720

GGGCAAAGCT AGAGAGGTAA GAAGAATCTA CAAATGTTCC TCGAGTTACA TGAACTTCCA 780

TCCCAATAAA CCCATTGGAA ACGAAAAATT TAAGTCAGAA GTGCATTTAA GGCTGGTCCG 840

AGTAGAATGA TTTTTACAAC GAATTGATCA CAACCAGTTA CAGATGTCTT TGTTCCTTCT 900

CCACTCCCAC TGCTTCACCT GACTAGCCTT TAAAAAAAAA A 941

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 843 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

CNAGGGATAA CCCCAAAGNT GGGAAATAAA CCCTCAATTA AAGGGGGAAC CAAAAAGCTG 60

GGAAGTTCCC CCCCGCGGTG GCGGCCNGNT CTAGGAACTA GTGGAATCCC CCGGGGCTGC 120

AGGGAATTCG GCACGGAGTG GGAATGTTGT TTGTATGATA CTATTTCCAC AAWATGCATT 180

GAGACTTGGT KTGTGGCCTA GGACATGGTC AATTCTTTYT AAATATTCCG TGAATTTCTT 240

TAGTGCATAT TCTCCGATGG GGGCTGTGGG GACAGAGTTC TAAATATGCC CATTAGATTA 300

AATCTCTTCA TTCTGTTGCT CACATCTTCT ATATCCTTAT TAATCTGTCA ATCTCTTCAA 360

GAGAGGTGTT ATTAAAATCT CTCACTGTAT GTGTCACTTT GCCCTTAAAA TTCTGATGAT 420

TTGCTTTATA AATGGTTATA ACCATTTTCC AGGAAGAACA TTAAAGAACT TTCCATTGGC 480

ATTATCCAGT TTCCCTCAAA ATACTGGTTT TTTTTATTTT GGCTNCTAAG CAGCTATGAA 540

TCCAGTTTCT CAGAAGCCCT TGTCTCAAGG CATTTGTTTC CAGATTACCT TGTTAGCATC 600

CACACTATGG GCTATTTTAG AAAAACAAAA AAAGTATCAA AATCATATAG CTATGATTTT 660

CCTGTGCTTG AAGGAGCCTT AAAGCTCATC TAGTCCAGCC AGTATTTGTT CATCCAAATT 720

CTGCCAAGAA ATCTCTATTG TCAAGATATT CTTTACCATC TTTGGGACAT TCTCATTATT 780

AGAAACAAAT CCTAAGAAGA AATTCTGCCA TAKACAACCC ATCCGTTCTT TAAAAAAAAA 840

AAA 843

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1018 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CTGTAATTTT TAATTTTCAT ATACCGTGCT TTGATTCTAA TTTTATTTTT TGAGTTCTCT 60

GAAGGTTACA TATACAGAGT GCTTCAGGAA TGATCATTTT GTTATTATTC ATGCTTCTTA 120

ACAATGTTGT TTTAGTCCAA GAAGATAATT GCCAGAGAAA GAATACAGTG CAGGAAAGAA 180

GARGCTGGAG CCAGTGGTGA AGARGGATTG AGARGACAGA CATTGTGGGA ATGAAATCAT 240

GAATAATCGT GTTTTTGAAT TGTCCAAAAA CTTCTACAAA CCATGAAATG TTGGAGTTTA 300

AATCTAATTG TTGAAAAATT CCCCACATTC CTTGTATCCC TTAGGTTGAG CATAATTCCA 360

CATCCGTGGA CTGATGCACT TCCCAAGAGG GGGCCTCATT AACTCTTCCG AGGCAGCAGC 420

AGCAAGGGCA CCCCCTCCTT TCCCCCCACA CCCCAYTTCT CATGGCTCTT CTTTCTCTCA 480

TCTCATGCTT AGGTTAGAAA AGGGCACAAG GTAAGGAAGC CCTTGGGAAT AGGCTGAATC 540

TGGCTATCTA ATTTGGTGCC AAATACTTAA TGTGCTTGAA TTTAAAAACA GCAAACATGT 600

AGAAAGGTAA TTATAATTAT GAGGCCAGTT CTTTAAGCTA GCTTTTTTTC CCCTCTCAAA 660

CAGCATATTG GCTTGGATGT CAGCAGGAGA AAGTGTTTTT TGCAATACAC ATAATGCATA 720

TATGGTCCTG TTAGCAATCT ATAGAAAATA GATATTGCTC ATTAAGGTAA ATATTTTTGT 780

TGATGAATGA TCTGGAATGG TCTGGACTTG TTGTGTGAAC AGGAAATTGC TCTGTAGGCT 840

TTGACTTGTG AGGTAAAGAG TGAGGCTGGT AAGATTAATT AAAGTAAATA CTGTGACAAT 900

AGGATGTCAA AACCAAAAAC GTGTTTCTGA AACTCAAGGA ATTAATGACA CATAGGGAAG 960

TTTTTGCCAT ATTAAGCATA GAGTAGGAGA GGCAAGTCAA GAATAAAAAA AAAAAAAA 1018

(2) -INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 661 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

TTTAAGAAAT TAGTGAATCC CCGGNTGCAG GGAATTCGGC ACGAGGAGGA GGCCGTCAGC 60

TGGCAGGAGC GCAGGATGGC AGCTGYTCCC CCGGGTTGCA CCCCCCCAGY TCTGCTGGAC 120

ATAAGYTGGT TAACAGAGAG CCTGGGAGCT GGGCAGCCTG TACCTGTGGA GTGCCGGCAC 180

CGCCTGGAGG TGGCTGGGCC AAGGAAGGGG CCTCTGAGCC CAGCATGGAT GCCTGCCTAT 240

GCCTGCCAGC GCCCTACGCC CCTCACACAC CACAACACTG GCCTMTCCGA GCTGCTGGAG 300

CATGGAGTGT GTGAGGAGGT GGAGAGAGTT CGGCGCTCAG AGAGGTACCA GACCATGAAG 360

GTGCGCAGGG CAGGGCTCGG ACCTACCCCA GGAATGTCCT GCCCTGGGAA TGACAACACA 420

GTCCACACCA TGCACGGGGA GGCAAACAGG GGCAGCTGAC CCAGCCCAGG GGTCAGANGA 480

GGTCTTGCCG AGGAAGTGGC AGCTAAGCTG ATACCTGATA TGCACWAGKC AGCCARGYGG 540

AGACAGGCAA GGAAGAAGCT TGTTTTGAGG ACAGAATTTT CTAGATCACT CAGCACCATC 600

TGGCTTTTGG GGCTTTTTGT TTTATTTTGT TTTTGAGACG GGGTCTCGCT CTGTCGCCCA 660

N 661

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 553 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

GGCACAGGGC TATTTGCCCC TCTCTCCACA TGACAGAACT GCTCTAAGTT TCTTTGCTGC 60

TCTTCTCAGC TGTCAGACGG CTTGCTGCTT GTTTTCCACA CCACCATGTC TATTCTTTGC 120

TGTCCTTWAC TCTGCCTGTT TTTTTCCTTT TGTATTTCTT CTGGCTCTTG TCCCTTTTCC 180

CACGTGTCWC AGCTTTCCTT TATTGCCACT TTCAGTCAGA GCAGTCCTGT GCTTCTGGTG 240

CCGGCATACA ATACTTACTT GAGTTTCTTG GCTTTTCTTG ACTGTGCATC TCTTACTTCA 300

ACATAGGAAT AGCCTGTCAT AGAATTTCTC CAGTTCCAGG GCTCAAGAGG GAGAGTGCCA 360

GAAAATTGAG ACTGTTTTCC CTGTCTTGGA TTGAATTCAT AAAGCAAAAC CAGTGTTTGT 420

GTGAGGGTTT GCTGTGTCAT GCCTATAGGT TGTTTGGGTG CAAACCTATA GAATCCAGCC 480

TGCGAAAAGA AAGRAACCAG AGAATANCAG CATCAGAACA ATGCTTGACA TCATTTCTCA 540

ATCAAGCAGT CCA 553

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 869 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GGCACGAGCT GCCAACACTG AGGTCTTCGT GGCTTCTCAC ATCTAGATGT ATCCCTCTCA 60

AATCTATCCT CTATCCAGGC ACCAGATTGA GGTATCTAAA ATGTCAACTT TCCAGTTACT 120

CCTTCTTATA CTAGCCCAAT CAACTTACAA GATAAAGTCC AAGCCCCTTC ATATGACAAA 180

CCACACCCTG CTTAACTCTC CAGGTTTGAA TCCTTCATCT CCTACTTTAA ACTTTAAAAC 240

CCAGCAGCAC GAAAGTGTCT CCTATGCATG TTGCCATATG CGTTCTCTCC ATCATGCATT 300

TGCCTGAGCA AGATGTCTTG AGTTAACATC TTATTCTTTA AGACTCATTG TGGTGGTAGA 360

CAGCCTTTAA TAACGGATCC TTGGCCAGGC ACAGTGACTC ACACCTGTAA TCCCAGAACT 420

TTGAAAGGCC AAAGAAGGAA GAAAGCTTGA GGCCAGTAGT TTGAGACCAG CCTGGGAAAC 480

AGAGAGATAT CCCATCTGTA CCAAAAATTT AAAAAAATAT TAGCAGGGAG TAGTGGCATG 540

CACAAGTGGT CCCAGCTCCA TGGGAGASTG AGGTAGGAAC ATCACTTGAG CCCAGGAAGT 600

CAAGGCTGCA GTGAACCATG ATCAGAACAT TGCANTCCAG CTTGGGTAAC AGAGTGAGAC 660

CTTAGGTCAG AAAAATGAAT AAATAAGCAT AAAATTTTAA AAACTTAGCC AGGCATGGTG 720

GCACACATCT GTGGTCCCTG CTACTTAGGA GGCTGAGGTG AGAGGATCCT TGAGCCCAGG 780

AGGTCAACAC TACAGTGAGC TATGATTGTG CCACTAAACT CCAACCTGGG TGAAAAAGCA 840

AAACCCTGCC AAAAAAAAAA AAAAAAACT 869

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 959 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

GGCGAGCCGA GATCGTGCCA TTGCACTCCA GCCTGGGCAA CAAGAGTGAA ACTCTGTCTC 60

AAAAAAAAAA AATTATAATA CTATATGCCA TAAAATGACA TTTCATATTT AAAGAGTTTT 120

TTAAAACTCT TGTATTCACA TGCCATAATT TGAAACCCTA TTTCACTGAA TGAGAATGGT 180

ATCTGTTGTC CTCATTTTTT CATTTTTATC CTTAACAATT TCCACCACAG CCAGTGCATA 240

TAATGGCAAT GACACCCAGG GATGGAATGA TAAGTTCCAT CRCMGCTCAG TCAAGACGCA 300

GACTTGATGT GGCCCCAACA ACAGTCAATA ATGGAGTCTC CAAAATAAAG CTCTATAGGA 360

AAGGTAAATA CCCGCTGCAC AAGAAACCAC AGCATCTAGG TTCTAACCCC ATCTCTATGA 420

AGAGCTTGCT GGGAGAGTTT TGACATTWAA CAATCTGTCT GATKGCCAAT TTTYTTCTTC 480

TATAAAATGA TAATGTTKGA YTCAAAGATC CAAAGTCAAT TCATGGTCTA AAACTTAATG 540

ATTTTTTTAG GTTTTGKGAC ATTTCACTGT ACACTGTAGT AATTTATATC TTATTTTCCC 600

ACTAATTTAG AAAAATATYT AAATGATCCT TAATTGGCAA TGGGTCCTAA GAATTTTGTT 660

TTAAATCCCT GTTACCCAAA AGAGCCCTTT TTTGTATCTC GCAGTAGTTA CAAGGATCTT 720

TCTAAATCTT AAAAAAAAAA AAAAAAGAAA GAAAGAAAAG AAAAGAAAAA AAGTCAGCCG 780

GGCGTGGTGG CTCATGCCTG TAATCCCAGC ACTTTGGGAC CAAGGTGGAC AGATCACGAG 840

GTCAGGAGAT GGAGACCATC CCGGCCAACA TGGAGAAACC CTGTCTCTAC TAAAAAAAAA 900

AAAAACTCGA GGGGGGCCCG GTACCCAATN CGCCGGCTAG TGGTCGTAAA ACAATCAAA 959

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1446 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

CGGGGCAGGG CTGTGTGGCA CCGCCAGGGA GCGGGCCCAC CTGAGTCACT TTATTGGGTT 60

CAGTCAACAC TTTCTTGCTC CCTGTTTTCT CTTCTGTGGG ATGATCTCAG ATGCAGGGGC 120

TGGTTTTGGG GTTTTCCTGC TTGTGCCAAG GGCTGGACAC TGCTGGGGGG CTGGAAAGCC 180

CCTCCCTTCC TGTCCTTCTG TGGCCTCCAT CCCCTCATGG GTGCTGCCAT CCTTCCTGGA 240

GAGAGGGAGG TGAAAGCTGG TGTGAGCCCA GTGGGTTCCC GCCCACTCAC CCAGGAGCTG 300

GCTGGGCCAG GACCGGGAGA GGGAGCACTG CTGCCCTCCT GGCCCTGCTC CTTCCGCAGT 360

TAGGGGTGGA CCGAGCCTCG CTTTCCCCAC TGTTCTGGAG GGAAGGGGAA GGAGGGGGTC 420

TTCAGGCTGG AGCCAGGCTG GGGGTGCTGG GTGGAGAGAT GAGATTTAGG GGGTGCCTCA 480

TGGGGTGGGC AGGCCTGGGG TGAAATRAGA AAGGCCCAGA ACGTGCAGGT CTGCGGAGGG 540

GAAGTGTCCT GAGTGAAGGA GGGGACCCCC ATCCTGGGGG ATGCTGGGAG TGAGTGAGTG 600

AGATGGCTGA GTGAGGGTTA TGGGGAGCCT GAGGTTTTAT GGGCCTGTGT ATCCCCTTCT 660

CCCGGCCCCA GCCTGCCTCC CTCCTGCCCG CCTGGCCCAC AGGTCTCCCT CTGGTCCCTG 720

TCCCTCTGGT GGTTGGGGAT GGAGCGGCAG CAAGGGGTGT AATGGGGCTG GGTTCTGTCT 780

TCTACAGGCC ACCCCGAGGT CCTCAGTGGT TGCCTGGGGA GCCGGACGGG GCTCCTGAGG 840

GGTACAGGTT GGGTGGGCCC TCCCTGAGGG TCTGGGGTCA GGCTTTGGCT CTGCTGCCTC 900

TCAGTCACCA AGTCACCTCC CTCTGAAAAT CCAGTCCCTT CTTTGGATGT CCTTGTGAGT 960

CACTCTGGGC CTGGCTGTCG TCCCTCCTCA GCTTCTTGTT CCTGGGACAA GGGTCAAGCC 1020

AGGATGGGCC CAGGCCTGGG ATCCCCCACC CCAGGACCCC CAGGCCCCCT CCCCTGCTGC 1080

TTTGCGGGGG GCAGGGCAGA AATGGACTCC TTTTGGGTCC CCGAGGTGGG GTCCCCTCCC 1140

AGCCCTGCAT CCTCCGTGCC STAGACCTGC TCCCCAGAGG AGGGGCCTTG ACCCACAGGA 1200

CGTGTGGTGG CGCCTGGCAC TCAGGGACCC CCAGCTGCCC CAGCCCTGGT CTCTGGCGCA 1260

TCTCTTCCCT CTTGTCCCGA AGATCTGCGC CTCTAGTGCC TTTTGAGGGG TTCCCATCAT 1320

CCCTCCCTGA TATTGTATTG AAAATATTAT GCACACTGTT CATGCTTCTA CTAATCAATA 1380

AACGCTTTAT TTAAAGCCAA AAAAAAAAAA AAAAAACTCG AGGGGGGGCC CGTACCCAAT 1440

TCGCCA 1446

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1471 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

CAAAAAATAA TAATGATAAT TTAAAATAAA TAAGTAACTA ATAAAAAGAT TTTATATCCC 60

AGTCTTATGA TGTTGGTTGG CAAGGCTAGA TAAAAAGATG TTAGAATGAA AGAACATATT 120

TTTAGTGATA TGTAAATGAA GGATTCTACA ATAGTCATAT ATTTTTATAT GAATGAATGT 180

TGGGTTGGGC TGGAGAGGTA TGTGTGTGTA AATATAAAGG TCTCACATTC AGAGTATAGC 240

TCTGAAATAA TGGAACTCAT GTCTACAATT CAACATGCAT CTGTATAGTT ACATCTCATG 300

TAAATATACA CAGACATATT TTGCAGCCAG TAATTGACAG TTAATGTCCA AAACAGGTGA 360

TTGATAGGTA ACAGAAATTA GATAACCACC AATTTTGCCC AAGAGAAAGA CTAGAAGGAC 420

TAAAAGCAGT TGAATGTATG GTACTGACAT TGTCATAAGC AGTCTGATAA CCAGTTTATT 480

GAAACGTGTG CATTAACAGA GAATTTAATT TTAAACCCAT AATTTCTCCT ATCCATTAAA 540

ATATTATAAT TGTTAGTAGT ATGAAACCAA CAGGAAATGT TTTTTAATCA TTTAGTGAGG 600

TGATTCATTT GTTTCATGGG CAAACACTAT CCAGGAAAAG CCTTGCTTGC CTGTTTCCCA 660

AAGAGCTCTA AGAAATAGAA TCAAGTGTAA AATGGTTCAG ACCATTCAGG ATTTCTTGTC 720

ACTCTTCTCA ACCCCGATCT TCCTGTTATT ACTGATGTTT GAAACCCTGT CATTAGCCCC 780

GGCCTGGTTA AAGCCCCTCA GAGTCACCTC TCATTCATAG CAATAGAATT CAACCCCAAG 840

TGGTTGATGG TGTCCCCAGC ACAGCCGAGA GACCTGATCT CTGGATTCAG TGCTTTTAGC 900

TCTTCGAGTT TACCCTAAGA TACCTTCGGG CAATATTTTT AACCAACCCA AAAGCTCTTC 960

AGGTCATTTC TGAAGAGGAC AAGGTGAATC TTGGCTTGGA ACACCATTTT TGGGCTCTTG 1020

CTACTGAATG AATCAGAAAG GAATTTTTTC TGAAGAGCAT TAGAAAGTAA AGGAGATGTT 1080

AAAATAAGTT CTTGAAGTAT GTTTTATATT TATCTAAAAC ACTGATTTTA AAAGTTTACA 1140

TTCAAATGTG TATTCAAAAG AAGTACTGAT TTGTAATTAT TATAGTTTGT GTGTATCATC 1200

CCCTTTTAAC CGTGCCTAAC AACTGTACTT AAATTTTGTT TTCCTAGTGT AACAAATGTT 1260

TCCCATAAGA TTTTCTAGAG CCAAATAATG GGAGTGAAAA ATTCCTTAAG TGTTATATAA 1320

GAAAATATAT TAGAAAATCA GCTTTGGATT ATACGATTTC TAAAATATAC TAATACAGAA 1380

TCCTCAGTAA TATGTTTTGA ATTGGATTTT TTCTCAGAAC TGTTACATAA TAAATAATAC 1440

ATCAACCAGA AAAAAAAAAA AAAAAAATTN C 1471

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1402 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

AGGGACGTCT TGCCTGAGGA GATGCCCATT TCTGTCCTGG RTTACCCTCA CTGCGTGGTG 60

CATGAGCTGC CAGAGCTGAC GGCGGAGAGT TTGGAAGCAG GTGACAGTAA CCAATTTTGC 120

TGGAGGAACC TCTTTTCTTG TATCAATCTG CTTCGGATCT TGAACAAGCT GACAAAGTGG 180

AAGCATTCAA GGACAATGAT GCTGGTGGTG TTCAAGTCAG CCCCCATCTT GAAGCGGGCC 240

CTAAAGGTGA AACAAGCCAT GATGCAGCTC TATGTGCTGA AGCTGCTCAA GGTACAGACC 300

AAATACTTGG GGCGGCAGTG GCGAAAGAGC AACATGAAGA CCATGTCTGC CATCTACCAG 360

AAGGTGCGGC ATCGGCTGAA CGACGACTGG GCATACGGCA ATGATCTTGA TGCCCGGCCT 420

TGGGACTTCC AGGCAGAGGA GTGTGCCCTT CGTGCCAACA TTGAACGCTT CAACGCCCGG 480

CGCTATGACC GGGCCCACAG CAACCCTGAC TTCCTGCCAG TGGACAACTG CCTGCAGAGT 540

GTCCTGGGCC AACGGGTGGA CCTCCCTGAG GACTTTCAGA TGAACTATGA CCTCTGGTTA 600

GAAAGGGAGG TCTTCTCCAA GCCCATTTCC TGGGAAGAGC TGCTGCAGTG AGGCTGTTGG 660

TTAGGGGACT GAAATGGAGA GAAAAGATGA TCTGAAGGTA CCTGTGGGAC TGTCCTAGTT 720

CATTGCTGCA GTGCTCCCAT CCCCCACCAG GTGGCAGCAC AGCCCCACTG TGTCTTCCGC 780

AGTCTGTCCT GGGCTTGGGT GAGCCCAGCT TGACCTCCCC TTGGTTCCCA GGGTCCTGCT 840

CCGAAGCAGT CATCTCTGCC TGAGATCCAT TCTTCCTTTA MTTCCCCCAM CCTCCTCTCT 900

TGGATATGGT TGGTTTTGGC TCATTTCACA ATCAGCCCAA GGYTGGGAAA GCTGGAATGG 960

GATGGGAACC CCTCCGCCGT GCATCTRAAT TTCAGGGGTC ATGCTGATGC CTCTCGAGAC 1020

ATACAAATCC TTGCCTTTGT CAGCTTGCAA AGGAGGAGAG TTTAGGATTA GGGCCAGGGC 1080

CAGAAAGTCG GTATCTTGGT TGTGCTCTGG GGTGGGGGTG GGGTGTTTCT GATGTTATTC 1140

CAGCCTCCTG CTACATTATA TCCAGAAGTA ATTGCGGAGG CTCCTTCAGC TGCCTCAGCA 1200

CTTTGATTTT GGACAGGGAC AAGGTAGGAA GAGAAGCTTC CCTTAACCAG AGGGGCCATT 1260

TTTCCTTTTG GCTTTCGAGG GCCTGTAAAT ATCTATATAT AATTCTGTGT GTATTCTGTG 1320

TCATGTTGGG GTTTTTAATG TGATTGTGTA TTCTGTTTAC ATTAAAAAGA AGCAAAAATA 1380

ATAAAAAAAA AAAAAAAAAA CT 1402

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1047 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

GGCACAGGGG ACTACAGGCA CCCACGACCA TACCCAGCTA ATTTTTGTAT TTTTTTGTAG 60

AGATGGGGTT TCACGATGTC GCCCAGGCTG GTCTTGAACT CCTGGGCTTG AGCGATCTTC 120

CCATCTTTCC ATCTTGGCCT CCTAAAGTGC TGGGACTGCA GGCATGAGCC ACCATGCCCA 180

GCCAAGATTC TTATTGATTA CCATGTTGCT TCAAGAAGCC AAGCCAGTTT CCAATATTCC 240

CCATTTGCTG GAGTCTTGGT ACTTTGGGTA GAAGCAACTG GTAAATTGTT AATTGGAACA 300

NTTGGTGGTG TAGATAACCA CGTATGGCCA AACCTAGAGC ATCTAGGCTC ACAATTACTA 360

TCCTGACTTG ATAACAAGTG TTCTGATATT AACCTGAAAA TGGGAATAAT GCCAAATCTG 420

TGTAACTTAA CATCTATATA CACAGTGGGG AGAACTGAAG TTATTAAACC TGGAATCTCT 480

GTGATCAAGG CTAACAGTAG TTATCTAAGA AGCAAAGGAC CTACAATTCT TAGACTTGGA 540

GTCATATTCT TTAAGGACGT GTTCTGAAAC TATATCAAGC ATCTGGTTTC CACGTATTTC 600

TCCCTCAGAA ATTATGAAGT ACAAGTAAAA ATGAAGGTAC AGGGTAAGAC ACATGCTGCT 660

TTCTTGCTCT TGAGTGGAGA CAGTTTTCCA GCCATCTTAA CCCCTTWACA CAAAACAATT 720

TGTGTTTTAT AGCAAATAAG TGACTCAACA TAATTTCAAT ATGATGTTTA TCCACCAGTA 780

CTTTCCTTTC AGCTTCTAGT CCCATAARTG GTTTGTGAAG TCATCGGTTA CATTAGCCAA 840

GATAGGCCTA GACTTGAAGT CTAGAATGTT TTTCCCACTA TATGCCAAAG TAGAATGTGG 900

GTATCTCAGG GTCATTTTTG TTGTTCAATT TCCCACCTGT ACAGTTGTTA TGATTCACTT 960

TCCTTATGTG TCTAATAAAT CTTGTTCCAT GAAATGATCA AAAAAAAAAA AAAAAAACT 1020

CGAGGGGGGG CCCGGTACCC AAATCGC 1047

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 990 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

TTGGAAAGGG TCTAGCTCTT TCTCATTCAC CAACTATATT AGAAGCACTT GAGGGAAATT 60

TACCACTCCA AATCCAAAGC AATGAACAGT CTTTTCTGGA TGATTTTATT GCCTGTGTCC 120

CAGGATCAAG TGGTGGAAGG CTTGCAAGGT GGCTTCAGCC AGATTCATAT GCGGATCCTC 180

AGAAAACATC TTTGATCCTG GAATAAGGAT GATATTCGTT GTGGTTGGCC TACCACCATA 240

ACTGTTCAAA CAAAAGACCA GTATGGGGAT GTGGTACATG TTCCCAATAT GAAGGTAATT 300

ATAACTGGAT TAAATTAGCA GACATCTATA TACTGGCTGC AATGACTGAT AAAATTTTAG 360

AAATGCCAAG TGCTGAGRGT CCATTTGTTC TACCCTCTTT ATATAAAGGG TGATGCTGAA 420

AGTTTGTTTA AATGACTTGT TTATATTAAT TAGTCCCCAA GTGTCCAAGT TACACCTGTT 480

TTTTTTGTGA GTTTGTTCTT TACATTTTGC TACCTGTTAC GGGGACTCAA AGGAGGGATA 540

AGAAAGTATC CATCTAAAGA GTGCTAGACA CATACAGTGA AGCCCCTCAA TATGTATTGA 600

TTGAATAAAT GCATGAAAGA ATACATTTTT AAATTTTGTG TATAGTTTTG AAAGACTCAA 660

GTACGTTCTG TGTTTGGTAT TACTGAAACC ACATTTTAAA AATAACACTC ATTAAGTTAG 720

AAATATATGA GTTTAGATTG TAAAAGAATG AGGAATTGAA ATAGTTGTAT ACCATATTGA 780

TGAATATAGA GTTTTTAGGA TACCTCTTAC CTGAAATATT AATAATAATG TTTNCAGAGC 840

ATATTATACA TAATTATTTG TGATTTAATC TGTTAATATG AATATCTCAT TTAAAACTTT 900

TATTTCTGAA AAAATTATAT TGAATAAAAT TTTATATAGG CAGTCCCCAG CCCTTTCCTC 960

CTTCAAAGTT GTCTTATAGA GTGATTGGTT 990

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1208 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

TAATCGCTAC TATAGGGAAA GCTGGTCGCT GCAGGTACCG GTCCGGAATT CCGGGTCGAC 60

CCACGCGTCC GAGCGAAATG GCGCCTCCGG CCCCCGGCCC GGCCTCCGGC GGCTCCGGGG 120

AGGTAGACGA GCTGTTCGAC GTAAAGAACG CCTTCTACAT CGGCAGCTAC CAGCAGTGCA 180

TAAACGAGGC GCASGGGTGA AGCTRTCAAG CCCAGAGAGA GACGTGGAGA GGGACGTCTT 240

CCTGTATAGA GCGTACCTGG CGCAGAGGAA GTTCGGTGTG GTCCTGGATG AGATCAAGCC 300

CTCCTCGGCC CCTGAGCTCC AGGCCGTGCG CATGTTTGCT GACTACCTCG CCCACGAGAG 360

TCGGAGGGAC AGCATCGTGG CCGAGCTGGA CCGAGAGATG AGCAGGAGCK TGGACGTGAC 420

CAACACCACC TTCCTGCTCA TGGCCGCCTC CATCTATCTC CACGACCAGA ACCCGGATGC 480

CGCCCTGCGT GCGCTGCACC AGGGGGACAG CCTGGAGTGC ACAGCCATGA CAGTGCAGAT 540

CCTGCTGAAG CTGGACCGCC TGGACCTCGC CCGGAAGGAG CTGAAGAGAA TGCAGGACCT 600

GGACGAGGAT GCCACCCTCA CCCAGCTCGC CACTGCCTGG GTCAGCCTGG CCACGGGTGG 660

TGAGAAGCTG CAGGATGCCT ACTACATCTT CCAGGAGATG GCTGACAAGT GCTCGCCCAC 720

CCTGCTGCTG CTCAATGGGC AGGCGGCCTG CCACATGGCC CAGGGCCGCT GGGAGGCCGC 780

TGAGGGCCTG CTGCAGGAGG CGCTAGACAA GGATAGTGGC TACCCRGAGA CGCTGGTCAA 840

CCTCATCGTC CTGTCCCAGC ACCTKGGCAA GCCCCCTGAG GTGACAAACC GATACCTGTC 900

CCAGCTGAAG GATGCCCACA GGTCCCATCC CTTCATCAAG GAGTACCAGG CCAAGGAGAA 960

CGACTTTGAC AGGCTGGTGC TACAGTACGC TCCCAGCGCT GAGGCTGGCC CAGAGCTGTC 1020

AGGACCATGA AGCCAGGACA GAGGCCAGGA GCCAGCCCTG CAGCCCTCCC CACCCGGCAT 1080

CCACCTGCAT CCCTCTGGGG CAGGAGCCCA CCCCCAGCAC CCCCATCTGT TAATAAATAT 1140

CTCAACTCCA RGGTGTTCCA CCTGAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1200

AAAAAAAA 1208

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1922 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

GTGCTGCGCT ACTGAGCAGC GCCATGGAGG ACTCTGAAGC ACTGGGCTTC GAACACATGG 60

GCCTCGATCC CCGGCTCCTT CAGGCTGTCA CCGATCTGGG CTGGTCGCGA CCTACGCTGA 120

TCCAGGAGAA GGCCATCCCA CTGGCCCTAG AAGGGAAGGA CCTCCTGGCT CGGGCCCGCA 180

CGGGCTCCGG GAAGACGGCC GCTTATGCTA TTCCGATGCT GCAGCTGTTG CTCCATAGGA 240

AGGCGACAGG TCCGGTGGTA GAACAGGCAG TGAGAGGCCT TGTTCTTGTT CCTACCAAGG 300

AGCTGGCACG GCAAGCACAG TCCATGATTC AGCAGCTGGC TACCTACTGT GCTCGGGATG 360

TCCGAGTGGC CAATGTCTCA GCTGCTGAAG ACTCAGTCTC TCAGAGAGCT GTGCTGATGG 420

AGAAGCCAGA TGTGGTAGTA GGGACCCCAT CTCGCATATT AAGCCACTTG CAGCAAGACA 480

GCCTGAAACT TCGTGACTCC CTGGAGCTTT TGGTGGTGGA CGAAGCTGAC CTTCTTTTTT 540

CCTTTGGCTT TGAAGAAGAG CTCAAGAGTC TCCTCTGTCA CTTGCCCCGG ATTTACCAGG 600

CTTTTCTCAT GTCAGCTACT TTTAACGAGG ACGTACAAGC ACTCAAGGAG CTGATATTAC 660

ATAACCCGGT TACCCTTAAG TTACAGGAGT CCCAGCTGCC TGGGCCAGAC CAGTTACAGC 720

AGTTTCAGGT GGTCTGTGAG ACTGAGGAAG ACAAATTCCT CCTGCTGTAT GCCCTGCTCA 780

AGCTGTCATT GATTCGGGGC AAGTCTCTGC TCTTTGTCAA CACTCTAGAA CGGAGTTACC 840

GGCTACGCCT GTTCTTGGAA CAGTTCAGCA TCCCCACCTG TGTGCTCAAT GGAGAGCTTC 900

CACTGCGCTC CAGGTGCCAC ATCATCTCAC AGTTCAACCA AGGCTTCTAC GACTGTGTCA 960

TAGCAACTGA TGCTGAAGTC CTGGGGGCCC CAGTCAAGGG CAAGCGTCGG GGCCGAGGGC 1020

CNAAAGGGGA CAAGGCCTCT GATCCGGAAG CAGGTGTGGC CCGGGGCATA GACTTCCACC 1080

ATGTGTCTGC TGTGCTCAAC TTTGATCTTC CCCCAACCCC TGAGGCCTAC ATCCATCGAG 1140

CTGGCAGGAC AGCACGCGCT AACAACCCAG GCATAGTCTT AACCTTTGTG CTTCCCACGG 1200

AGCAGTTCCA CTTAGGCAAG ATTGAGGAGC TTCTCAGTGG AGAGAACAGG GGCCCCATTC 1260

TGCTCCCCTA CCAGTTCCGG ATGGAGGAGA TCGAGGGCTT CCGCTATCGC TGCAGGGATG 1320

CCATGCGCTC AGTGACTAAG CAGGCCATTC GGGAGGCAAG ATTGAAGGAG ATCAAGGAAG 1380

AGCTTCTGCA TTCTGAGAAG CTTAAGACAT ACTTTGAAGA CAACCCTAGG GACCTCCAGC 1440

TGCTGCGGCA TGACCTACCT TTGCACCCCG CAGTGGTGAA GCCCCACCTG GGCCATGTTC 1500

CTGACTACCT GGTTCCTCCT GCTCTCCGTG GCCTGGTRCG CCCTCACAAG AAGCGGAAGA 1560

AGCTGTCTTC CTCTTGTAGG AAGGCCAAGA GAGCAAAGTC CCAGAACCCA CTGCGCAGCT 1620

TCAAGCACAA AGGAAAGAAA TTCAGACCCA CAGCCAAGCC CTCCTGAGGT TGTTGGGCCT 1680

CTCTGGAGCT GAGCACATTG TGGAGCACAG GCTTACACCC TTCGTGGACA GGCGAGGCTC 1740

TGGTGCTTAC TGCACAGCCT GAACAGACAG TTCTGGGGCC GGCAGTGCTG GGCCCTTTAG 1800

CTCCTTGGCA CTTCCAAGCT GGCATCTTGC CCCTTGACAA CAGAATAAAA ATTTTAGCTG 1860

CCCCAAAAAA AAAAAAAAAA AAAAAAACTC GAGGGGGGGC CCGTACCCAA TTCGCCCTAT 1920

AA 1922

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1951 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

TCGTCCCCAG AGCGGGCTGA GCCCCAGGCG SAGGGTGGCG GGGGAGCCTG GGGGAGCCGC 60

CGCCACCTCC ACGGGCCTCT CTGAGCTCGG ACACCAGCGC CCTGTCCTAT GACTCTGTCA 120

AGTACACGCT GGTGGTAGAT GAGCATGCAC AGCTGGAGCT GGTGAGCCTG CGCCGTGCTT 180

CGGAGACTAC AGTGACGAGA GTGACTCTGC CACCGTCTAT GACAACTGTG CCTCCGTCTC 240

CTCGCCCTAT GAGTCGGCCA TCGGAGAGGA ATATGAGGAG GCCCCGCGGC CCCAGCCCCC 300

TGCCTGCCTC TCCGAGGAAC TCCACGCCTG ATGAACCCGA CGTCCATTTC TCCAAGAAAT 360

TCCTGAACGT YTTCATGAGT GGCCGCTCCC GCTCCTCCAG TGCTGAGTCC TTCGGGCTGT 420

TCTCCTGCAT CATCAACGGG GAGGAGCAGG AGCAGACCCA CCGGGCCATA TTCAGGTTTG 480

TGCCTCGACA CGAAGACGAA CTTGAGCTGG AAGTGGATGA CCCTCTGCTA GTGGAGCTCC 540

AGGCTGAAGA CTACTGGTAC GAGGCCTACA ACATGCGCAC TGGTGCCCGG GGTGTCTTTC 600

CTGCCTATTA CGCCATCGAG GTCACCAAGG AGCCCGAGCA CATGGCAGCC CTGGCCAAAA 660

ACAGTGACTG GGTGGACCAG TTCCGGGTGA AGTTCCTGGG CTCAGTCCAG GTTCCCTATC 720

ACAAGGGCAA TGACGTCCTC TGTGCTGCTA TGCAAAAGAT TGCCACCACC CGCCGGCTCA 780

CCGTGCACTT TAACCCGCCC TCCAGCTGTG TCCTGGAGAT CAGCGTGCGG GGTGTGAAGA 840

TAGGCGTCAA GGCCGATGAC TCCCAGGAGG CCAAGGGGAA TAAATGTAGC CACTTTTTCC 900

AGTTAAAAAA CATCTCTTTC TGCGGATATC ATCCAAAGAA CAACAAGTAC TTTGGGTTCA 960

TCACCAAGCA CCCCGCCGAC CACCGGTTTG CCTGCCACGT CTTTGTGTCT GAAGACTCCA 1020

CCAAAGCCCT GGCAGAGTCC GTGGGGAGAG CATTCCAGCA GTTCTACAAG CAGTTTGTGG 1080

AGTACACCTG CCCCACAGAA GATATCTACC TGGAGTAGCT GTGCAGCCCC GCCCTCTGCG 1140

TCCCCCAGCC CTCAGGCCAG TGCCAGGACA GCTGGCTGCT GACAGGATGT GGCACTGCTT 1200

GAGGAGGGGC ACCTGCCACC GCCAGAGGAC AAGGAAGTGG GGCGCTGGCC CAGGGTAGGG 1260

GAGGGTGGGG CAATGGGGAG AGGCAAATGC AGTTTATTGT AATATATGGG ATTAGATTCA 1320

TCTATGGAGG GCAGAGTGGG CTGCCTGGGG ATTGGGAGGG ACAGGGCTTG GGGAGCAGGT 1380

CTCTGGCAGA GAAGGATGTC CGTTCCAGGA GCACACGGCC CTGCCCCATC CTGGGCCTTA 1440

CCTCCCCTGC CAGGGCTCGG GCGCTGTGGC TCCTGCCTTG ATGAAGCCCG TGTCCTGCCT 1500

TGATGAAGCC TGTGCCACCT GCAAGTGCCC GCCCTGCCCC TGCCCCAACC CCCACCGAAG 1560

AGCCCTGAGC TCAGGCTGAG CCCAGCCACC TCCCAAGGAC TTTCCAGTGA GGAAATGGCA 1620

ACACGTGGAG GTGAAGTCCC TGTTCTCAGC TCCGTCATCT GCGGGGCTTC TGGGTGGCTC 1680

CTGCCACTGA CCTCACCGGC ATGCTGGCCT GTGGCAGGCC TAGGACCTCA GGCGGGGAGG 1740

AGGAGCTGCC GCAAGGCCCT GTCCCAGCAG AAGAGGGAGG CTTCCTGACT GACACAGGCC 1800

AGCCCCATCT TGGTCCTGTC ACCCTGGCCC CAACTATTAA AGTGCCATTT CCTGTCAAAA 1860

AAAAAAAAAA AAAATCGGGG GGGGCCCGGA ANCCAATTTC CCCCAAAAAG GGGGGTTATA 1920

AAAATTCCCN GGCNGTGTTT TTAAAAATTC G 1951

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3989 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

GGCACAGGCC GCAGGGNACC TATGGGCGCA TATAGGTTGT AATGAAACTG TAGTCTCAGT 60

TGGAAGCCTA GACATGAAAT GGGTCAGTGA GCAAGGCTCT ATTCCTAGTC TCCAGCCATG 120

CCTGTGGAAC CTGARCCCRC TCTCAGCACA TTGGACCCAG GCAGATGYAA AAAATTCACA 180

GAACTATGAT TTGGACTCAA GGGTTTGTAG ATTTCCTCCT TCATTCTAAT TTCAGTGTCT 240

AAAATTCTTG CATCCRTGAA CGAGCTGGGC ATTTGATGAG ACAGGGCYGA ATACTGCAGT 300

TTTCCTCCTA GAAATCATCT GGGGCATTTT CTTTGAACTG ATGGGAACAA TAAGGCATAA 360

CTGTTTGCAC AAACTTGGGA TAARTGATTT TGGGATAACG ATCTACCAGA ATGGGGATAT 420

TTCACCCTTG GTTCTGAGAT GCAAACCAAA GAATATCATG ACCAGCTTTC AGGCCTCCTG 480

AAGTATATCT CTCACATTGT CCTGTTCTCA TGCTGAGGAG CCTGAGATCC CTGTGTGGGG 540

ATTAGACAGT GGACTGTTAT GGGTGTAGGT GAATTGGCTT ATTTTGTCTG TCCCTGTCTG 600

AATGTATTGC AGGAAYTAAA AAGGACCAAG AAGAGGAAGA AGACCAAGGC CCACCATGCC 660

CCAGGCTCAG CAGGGAGCTG CTGGAGGTAG TAGAGCCTGA AGTCTTGCAG GACTCACTGG 720

ATAGATGTTA TTCAACTCCT TCCAGTTGTC TTGAACAGCC TGACTCCTGC CAGCCCTATG 780

GAAGTTCCTT TTATGCATTG GAGGAAAAAC ATGTTGGCTT TTCTCTTGAC GTGGGAGAAA 840

TTGAAAAGAA GGGGAAGGGG AAGAAAAGAA GGGGAAGAAG ATCAAAGAAG GAAAGAAGAA 900

GGGGAAGAAA AGAAGGGGAA GAAGATCAAA ACCCACCATG CCCCAGGCTC AGCAGGGAGC 960

TGCTGGATGA GAAAGRGCCT GAAGTCTTGC AGGACTCACT GGATAGATGT TATTCAACTC 1020

CTTCAGTTGT GTTGAACTGT GTGACTCATG CCAGCCCTAC AGAAGTGCCT TTTATGTATT 1080

GGAGCAACAG CATGTTGGCT TGGCTGTTGA CATGGATGAA ATTGAAAAGT ACCAAGAAGT 1140

GGAAGAAGAC CAAGACCCAT CATGCCCCAG GCTCAGCAGG GAGCTGCTGG ATGAGAAAGA 1200

GCCTGAAGTC TTGCAGGACT CACTGGATAG ATGTTATTCG ACTCCTTCAG GTTATCTTGA 1260

ACTGCCTGAC TTAGGCCAGC CCTACAGCAG TGCKGTTTAC TCATTGGAGG AMCAKTACCT 1320

TGGCTTKKCT CTTGACGTGG ASAAATTGAA AAGAAGGGGA AGGGGAARAA AAGAAGGGGA 1380

AGAAGATCAA AGAAGGAAAG AAGAAGGGGA AGAAAAGAAG GGGAAGAAGA TCAAAACCCA 1440

CCATGCCCCA GGCTCAGCAG GGAGCTGCTG GATGAGAAAG GGCCTGAAGT CTTGCAGGAC 1500

TCACTGGATA GATGTTATTC AACTCCTTCA GGTTGTCTTG AACTGACTGA CTCATGCCAG 1560

CCCTACAGAA GTGCCTTTTA YRTATTGGAG CAACAGYGTG TTGGCTTGGC TGTTGACATG 1620

GATGAAATTG AAAAGTACCA AGAAGTGGAA GAAGACCAAG ACCCATCATG CCCCAGGCTC 1680

AGCAGGGAGC TGCTGGATGA GAAAGAGCCT GAAGTCTTGC AGGACTCACT GGATAGATGT 1740

TATTCGACTC CTTCAGGTTA TCTTGAACTG CCTGACTTAG GCCAGCCCTA CAGCAGTGCT 1800

GTTTACTCAT TGGAGGAACA GTACCTTGGC TTGGCTCTTG ACGTGGACAG AATTAAAAAG 1860

GACCAAGAAG AGGAAGAAGA CCAAGGCCCA CCATGCCCCA GGCTCAGCAG GGAGCTGCTG 1920

GAGGTAGTAG AGCCTGAAGT CTTGCAGGAC TCACTGGATA GATGTTATTC AACTCCTTCC 1980

AGTTGTCTTG AACAGCCTGA CTCCTGCCAG CCCTATGGAA GTTCCTTTTA TGCATTGGAG 2040

GAAAAACATG TTGGCTTTTC TCTTGACGTG GGAGAAATTG AAAAGAAGGG GAAGGGGAAG 2100

AAAAGAAGGG GAAGAAGATC AAMGAAGRAA AGAAGAAGGG GAAGAAAAGA AGGGGAAGAA 2160

GATCAAAACC CACCATGCCC CAGGCTCAAC GGCGTGCTGA TGGAAGTGGA AGAGCSTGAA 2220

GTCTTACAGG ACTCACTGGA TAGATGTTAT TCGACTCCGT CAATGTACTT TGAACTACCT 2280

GACTCATTCC AGCACTACAG AAGTGTGTTT TACTCATTTG AGGAACAGCA CATCAGCTTC 2340

GCCCTTTACG TGGACAATAG GTTTTTTACT TTGACGGTGA CAAGTCTCCA CCTGGTGTTC 2400

CAGATGGGAG TCATATTCCC ACAATAAGCA GCCCTTASTA AKCCGAGAGA TGTCATTCCT 2460

GCAGGCAGGA CCTATAGGCA MGTGAAGATT TGAATGAAAG TACAGTTCCA TTTGGAAGCC 2520

CAGACATAGG ATGGGTCAGT GGGCATGGCT CTATTCCTAT TCTCAAACCA TGCCAGTGGC 2580

AACCTGTGCT CAGTCTGAAG ACAATGGACC CACGTTAGGT GTGACACGTT CACATAACTG 2640

TGCAGCACAT GCCGGGAGTG ATCAGTCRGA CATTTTAATT TGAACCACGT ATCTCTGGGT 2700

AGCTACAAAA TTCCTCAGGG ATTTCATTTT GCAGGCATGT CTCTGAGCTT CTATACCTGC 2760

TCAAGGTCAK TGTCATCTTT GTGTTTAGCT CATCCAAAGG TGTTACCCTG GTTTCAATGA 2820

ACCTAACCTC ATTCTTTGTG TCTTCAGTGT TGGCTTGTTT TAGCTGATCC ATCTGTAACA 2880

CAGGAGGGAT CCTTGGCTGA GGATTGTATT TCAGAACCAC CAACTGCTCT TGACAATTGT 2940

TAACCCGCTA GRCTCCTTTG GTTAGAGAAG CCACAGTCCT TCAGCCTCCA ATTGGTGTCA 3000

GTACTTAGGA AGACCACAGC TAGATGGACA AACAGCATTG GGAGGCCTTA GCCCTGCTCC 3060

TCTCRATTCC ATCCTGTAGA GAACAGGAGT CAGGAGCCGC TGGCAGGAGA CAGCATGTCA 3120

CCCAGGACTC TGCCGGTGCA GAATATGAAC AAYGCCATGT TCTTGCAGAA AACGCTTAGC 3180

CTGAGTTTCA TAGGAGGTAA TCACCAGACA ACTGCAGAAT GTRGARCACT GAGCAGGACA 3240

GCTGACCTGT CTCCTTCACA TAGTCCATRT CACCACAAAT CACACAACAA AAAGGAGARG 3300

AGATATTTTG GGTTCAAAAA AAGTAAAAAG ATAATGTAGC TGCATTTCTT TAGTTATTTT 3360

GARCCCCAAA TATTTCCTCA TCTTTTTGTT GTTGTCATKG ATGGTGGTGA CATGGACTTG 3420

TTTATAGAGG ACAGGTCAGC TGTCTGGCTC AGTGATCTAC ATTCTGAAGT TGTCTGAAAA 3480

TGTCTTCATG ATTAAATTCA GCCTAAACGT TTTGCCGGGA ACACTGCAGA GACAATGCTG 3540

TGAGTTTCCA ACCTYAGCCC ATCTGCGGGC AGAGAAGGTC TAGTTTGTCC ATCASCATTA 3600

TCATGATATC AGGACTGGTT ACTTGGTTAA GGAGGGGTCT AGGAGATCTG TCCCTTTTAG 3660

AGACACCTTA CTTATAATGA AGTATTTGGG AGGGTGGTTT TCAAAATTAG AAATGTCCTG 3720

TATTCCRATG ATCATCCTGT AAACATTTTA TCATTTATTA ATCATCCCTG CCTGTGTCTA 3780

TTATTATATT CATATCTCTA CGCTGGAAAC TTTCTGCCTC AATGTTTACT GTGCCTTTGT 3840

TTTTGCTAGT GTGTGTTGTT GAAAAAAAAA ACATTCTCTG CCTGAGTTTT AATTTTTGTC 3900

CAAAGTTATT TTAATCTATA CAATTAAAAG CTTTTGCCTA TCAAAAAAAA AAAAAAAAAA 3960

AAAAAAAAAA AAAAAGCGGA CGCGTGGGC 3989

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3735 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

CTGCTGTTCG CTGGCTGGGC TCCGCAGCAG GCTTGGCCAG CSGCTGACGG GTCGGCGGGC 60

GGGTTTGTGT GAACAGGCAC GCAGCTGCAG ATTTTATTCT GGTAGTGCAN CCCTCTCAAA 120

GGTTGAAGGA ACTGATGTAA CAGGGATTGA AGAAGTAGTA ATTCCAAAAA AGAAAACTTG 180

GGATAAAGTA GCCGTTCTTC AGGCACTTGC ATCCACAGTA AACAGGGATA CCACAGCTGT 240

GCCTTATGTG TTTCAAGATG ATCCTTACCT TATGCCAGCA TCATCTTTGG AATCTCGTTC 300

ATTTTTACTG GCAAAGAAAT CCGGGGAGAA TGTGGCCAAG TTTATTATTA ATTCATACCC 360

CAAATATTTT CAGAAGGACA TAGCTGAACC TCATATACCG TGTTTAATGC CTGAGTACTT 420

TGAACCTCAG ATCAAAGACA TAAGTGAAGC CGCCCTGAAG GAACGAATTG AGCTCAGAAA 480

AGTCAAAGCC TCTGTGGACA TGTTTGATCA GCTTTTGCAA GCAGGAACCA CTGTGTCTCT 540

TGAAACAACA AATAGTCTCT TGGATTTWTT GTGTTACTAT GGTGACCAGG AGCCCTCAAC 600

TGATTACCAT TTTCAACAAA CTGGACAGTC AGAAGCATTG GAAGAGGAAA ATGATGAGAC 660

ATCTAGGAGG AAAGCTGGTC ATCAGTTTGG AGTTACATGG CGAGCAAAAA ACAACGCTGA 720

GAGAATCTTT TCTCTAATGC CAGAGAAAAA TGAACATTCC TATTGCACAA TGATCCGAGG 780

AATGGTGAAG CACCGAGCTT ATGAGCAGGC ATTAAACTTG TACACTGAGT TACTAAACAA 840

CAGACTCCAT GCTGATGTAT ACACATTTAA TGCATTGATT GAAGCAACAG TATGTGCGAT 900

AAATGAGAAA TTTGAGGAAA AATGGAGTAA AATACTGGAG CTGCTAAGAC ACATGGTTGC 960

ACAGAAGGTG AAACCAAATC TTCAGACTTT TAATACCATT CTGAAATGTC TCCGAAGATT 1020

TCATGTGTTT GCAAGATCGC CAGCCTTACA GGTTTTACGT GAAATGAAAG CCATTGGAAT 1080

AGAACCCTCG CTTGCAACAT ATCACCATAT TATTCGCCTG TTTGATCAAC CTGGAGACCC 1140

TTTAAAGAGA TCATCCTTCA TCATTTATGA TATAATGAAT GAATTAATGG GAAAGAGATT 1200

TTCTCCAAAG GACCCGGATG ATGATAAGTT TTTTCAGTCA GCCATGAGCA TATGCTCATC 1260

TCTCAGAGAT CTAGAACTTG CCTACCAAGT ACATGGCCTT TTAAAAACCG GAGACAACTG 1320

GAAATTCATT GGACCTGATC AACATCGTAA TTTCTATTAT TCCAAGTTCT TCGATTTGAT 1380

TTGTCTAATG GAACAAATTG ATGTTACCTT GAAGTGGTAT GAGGACCTGA TACCTTCAGC 1440

CTACTTTCCC CACTCCCAAA CAATGATACA TCTTCTCCAA GCATTGGATG TGGCCAATCG 1500

GCTAGAAGTG ATTCCTAAAA TTTGGAAAGA TAGTAAAGAA TATGGTCATA CTTTCCGCAG 1560

TGACCTGAGA GAAGAGATCC TGATGCTCAT GGCAAGGGAC AAGCACCCAC CAGAGCTTCA 1620

GGTGGCATTT GCTGACTGTG CTGCTGATAT CAAATCTGCG TATGAAAGCC AACCCATCAG 1680

ACAGACTGCT CAGGATTGGC CAGCCACCTC TCTCAACTGT ATAGCTATCC TCTTTTTAAG 1740

GGCTGGGAGA ACTCAGGAAG CCTGGAAAAT GTTGGGGCTT TTCAGGAAGC ATAATAAGAT 1800

TCCTAGAAGT GAGTTGCTGA ATGAGCTTAT GGACAGTGCA AAAGTGTCTA ACAGCCCTTC 1860

CCAGGCCATT GAAGTAGTAG AGCTGGCAAG TGCCTTCAGC TTACCTATTT GTGAGGGCCT 1920

CACCCAGAGA GTAATGAGTG ATTTTGCAAT CAACCAGGAA CAAAAGGAAG CCCTAAGTAA 1980

TCTAACTGCA TTGACCAGTG ACAGTGATAC TGACAGCAGC AGTGACAGCG ACAGTGACAC 2040

CAGTGAAGGC AAATGAAAGT GGAGATTCAG GAGCAGCAAT GGTCTCACCA TAGCTGCTGG 2100

AATCACACCT GAGAACTGAG ATATACCAAT ATTTAACATT GTTACAAAGA AGAAAAGATA 2160

CAGATTTGGT GAATTTGTTA CTGTGAGGTA CAGTCAGTAC ACAGCTGACT TATGTAGATT 2220

TAAGCTGCTA ATATGCTACT TAACCATCTA TTAATGCACC ATTAAAGGCT TAGCATTTAA 2280

GTAGCAACAT TGCGGTTTTC AGACACATGG TGAGGTCCAT GGCTCTTGTC ATCAGGATAA 2340

GCCTGCACAC CTAGAGTGTC GGTGAGCTGA CCTCACGATG CTGTCCTCGT GCGATTGCCC 2400

TCTCCTGCTG CTGGACTTCT GCCTTTGTTG GCCTGATGTG CTGCTGTGAT GCTGGTCCTT 2460

CATCTTAGGT GTTCATGCAG TTCTAACACA GTTGGGGTTG GGTCAATAGT TTCCCAATTT 2520

CAGGATATTT CGATGTCAGA AATAACGCAT CTTAGGAATG ACTAAACAAG ATAATGGCAG 2580

TTTAGGCTGC ACAACTGGTA AAATGACTGT AGATAAATGT TGTAATTAGT GTACACGTTT 2640

GTATTTTTGT TAATATAGCC GCTGCCATAG TTTTCTAACT TGAACAGCCA TGAATGTTTC 2700

ATGTCTCCCT TTTTTTTTTG TCTATAGCTG TTACCTATTT TAGTGGTTGA AATGAGAGCT 2760

AGTGATGACA GAAGGATGTG GAATGTCTTC TTGACATCAT TGTGTATTGC TGGTAATCAA 2820

GTTGGTAACG ACTACTTCTA GCAGCTCTTA CCACTATGAC TTAAGTGGTC CTGGAAGGCA 2880

GTAAGTGGAG GTTTGCAGCA TTCCTGCCTT CATGAGGGCT TCTACCACTG ACCACTTTGC 2940

ACGTACCTGG CTCCCAGATT TACTTAGGTA CCCCACGAGT CGTCCACATA AGCAGCTTCA 3000

TCTTTACCTT GCCAGAGTTG ACAATTATGG GATACTCTAG TCTACTTATA CTTGTGTTCC 3060

CATCTGTCTG CCATCCTCTG AAGGCCAGGA CCCAGTCATA CATCCTTAGA AACCAAAGTA 3120

TGGTTTTTGT TTTCTCTTGG AATGTCAGGT CTTAAGGCAT TTAATTGAGG GACAAAAAAA 3180

AAAAAAAGCC GATATAGTAG CTAGCTACTT AAGCATCCAT GGGTATTGCT CCATATCAAA 3240

GCAGATTTGC AGGACAGAAA GAGTAAATTA GCCTTCAGTC TTGGTTTACA GCTTCCAAGG 3300

AGAGCCTTGG CCACCTGAAA TGTTAACTCG GTCCCTTCCT GTCTCTAGTT CATCAGCACC 3360

TGCAGATGCC TGACTCTTGT TAGCCTTACT ATTCAATACA GTCCTTAGAT TCACGGTATG 3420

CCTCTTCCTA TCCAGGCACC TATTCTGAAT CACCATGTTG CTCTGCAGCT AGAGTTGATA 3480

GGAGAAAATC CATTTGGGTA GATGGCCTAT GAATTTGTAG TAGACTTTCA AAATGAGTGA 3540

TTTGTTAGCT TGGTACTTTT AAGTTTGTGG TACAGATCCT CCAAACCCAT ACTCTGAGCA 3600

ATTAACTGCC TTGAACATAG AGAAAATTAA GGCCTCACAG GATGAGTCTC CATTCTCTGT 3660

AAATGCTTAT TTTATCATAG TCTTTAGCCN CTACTATGAG TAAAATGTTC TCTTCNGCCG 3720

GGTGTGGTGA CTCAC 3735

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1667 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

TAGTAATTCA TTTAACTCCT CTTACATGAG TAGCGACAAT GAGTCAGATA TCGAAGATGA 60

AGACTTAAAG TTAGAGCTGC GACGACTACG AGATAAACAT CTCAAAGAGA TTCAGGACCT 120

GCAGAGTCGC CAGAAGCATG AAATTGAATC TTTGTATACC AAACTGGGCA AGGTGCCCCC 180

TGCTGTTATT ATTCCCCCAG CTGCTCCCCT TTCAGGGAGA AGACGACGAC CCACTAAAAG 240

CAAAGGCAGC AAATCTAGTC GAAGCAGTTC CTTGGGGAAT AAAAGCCCCC AGCTTTCAGG 300

TAACCTGTCT GGTCAGAGTG CAGCTTCAGT CTTGCACCCC CAGCAGACCC TCCACCCTCC 360

TGGCAACATC CCAGAGTCCG GGCAGAATCA GCTGTTACAG CCCCTTAAGC CATCTCCCTC 420

CAGTGACAAC CTCTATTCAG CCTTCACCAG TGATGGTGCC ATTTCAGTAC CAAGCCTTTC 480

TGCTCCAGGT CAAGGAACCA GCAGCACAAA CACTGTTGGG GCAACAGTGA ACAGCCAAGC 540

CGCCCAAGCT CAGCCTCCTG CCATGACGTC CAGCAGGAAG GGCACATTCA CAGATGACTT 600

GCACAAGTTG GTAGACAATT GGGCCCGAGA TGCCATGAAT CTCTCAGGCA GGAGAGGAAG 660

CAAAGGGCAC ATGAATTATG AGGGCCCTGG AATGGCAAGG AAGTTCTCTG CACCTGGGCA 720

ACTGTGCATC TCCATGACCT CGAACCTGGG TGGCTCTGCC CCCATCTCTG CAGCATCAGC 780

TACCTCTCTA GGTCACTTCA CCAAGTCTAT GTGCCCCCCA CAGCAGTATG GCTTTCCAGC 840

TACCCCATTT GGCGCTCAAT GGAGTGGGAC GGGTGGCCCA GCACCACAGC CACTTGGCCA 900

GTTCCAACCT GTGGGAACTG CCTCCTTGCA GAATTTCAAC ATCAGCAATT TGCAGAAATC 960

CATCAGCAAC CCCCCAGGCT CCAACCTGCG GACCACTTAG ACCTAGAGAC ATTAACTGAA 1020

TAGATCTGGG GGCAGGAGAT GGAATGCTGA GGGGGTGGGT GGGGGTGGGA AGTAGCCTAT 1080

ATACTAACTA CTAGTGCTGC ATTTAACTGG TTATTTCTTG CCAGAGGGGA ATGTTTTTAA 1140

TACTGCATTG AGCCCTCAGA ATGGAGAGTC TCCCCCGCTC CAGTTATTGG AATGGGAGAG 1200

GAAGGAAAGA ACAGCTTTTT TGTCAAGGGG CAGCTTCAGA CCATGCTTTC CTGTTTATCT 1260

ATACTCAGTA ATGAGGATGA GGGCTAGGAA AGTCTTGTTC ATAAGGAAGC TGGAGAACTC 1320

AATGTAAAAT CAAACCCATC TGTAATTTCG AGTGGGTGGA GCTCTTGCTT TTGGTACATG 1380

CCCTGAATCC CTCACTCCCT CAAGAATCCG AACCACAGGA CAAAAACCAC CTACTGGGCT 1440

CTCTCCTACC CTGCCCTCCT CCCTTTTTTT TACCCCTCTC TTTTTTATTT TTTCTTTGCT 1500

CTTTAGAACC CAGTGAAAAA TACCAGGGTA CTGGGGTGCA ACTCTTTCTT ATGATAGGTC 1560

ATTAGTGCTT TAAGCAAAAG ATATTAGCAG CTTTGACTGC AGCATTAGCA ATTAGGRAAA 1620

AAAAAAANWA AAAACTCGAG GGGGGGCCCG GTTACCCAAT TCGCCCT 1667

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1408 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

ATTACACACC TGAGCACTGT GCCTGGCAAG ACCTGTCTTA ATAGATTAGA GAACCACTGA 60

TAGATGGTCA GCTTTCTGTA GCAGTGAGAA CCCTACATTT CAAATGTGGA TAGCACCTTT 120

GCGGGGAAAC ATCACTTGGC ACATCTGCAT TCTTTTTTGA CACAGGGTCT CACTCTGTTG 180

CCCAGGCTAG AGTGCATGGC ACGATCTTAG CTCACTGCAA CCTCCACCTC CCAAGTTCAA 240

GCGATTCTTC TGCCTCAGCC TCCTGAGCAG CTGGGATCAC AGACATGCGC TACCATGCCC 300

AGCTAATTTT TTGTATTTTT TGTKTGTTTG TTTTTGTTTK TAAGTAGAGA CGGGCTTTCA 360

CCACGTTGGS CAGGCAGGTC TCGAACTCCT GAMCTCAGGT GATCCACCCA CATCTGCGTT 420

CCAATATCTT TCTCAACATA ATGATAGCCG TAATTAATAT TTTCCAGTAC ATTTTTATGC 480

CTTTACACAC GAGAGTGGTA GACAGACACA AACCCAGATC TGTCTGACTC CAAAGCCCGT 540

TTGTCATCAT TCCTTTTACG GTATCCTATA GTGGTATCCT TTACAGAAAG ACAGCTTTTA 600

CCCAACAAAG ACTTAACTTC CCAGGATGCC AGAAGGACAA AGCGGGATTG CTTTTAAGRA 660

GRAAGTTATC AAGAMCTTAT TTTATAAATG AGATTAGATA GGGAAAGGCA ATTTATCTTT 720

ATTAAAAACT GAAAAGGCCA GCATAGGGAA GGAGGTCCTT CGGTGGTCTT TTTCAGGGAA 780

ATACTTCAGT TGCTTTTATT AGAAACAGAT AGTACCTAAG GTTTTGAGGT AGGWACAGCT 840

TAAGGCATGC TAATGKTCAT GGGTCCTTCC ATAGTCATTT TKGTATTTTG GTTWACATTT 900

GAGCAATAGG CAGCCCTTCA CTGCTGCTGG AYTCATTCCT GCCAYTATTA CAGGTGACAG 960

AGGAGACAGG AGGTATGTCT TTTCTATTTT TAWACATGCT TTATATTTAA CACAAGCTCT 1020

TGGGTATCTT AGATAAACAG AAGTTGCCTA GCACTCCTTT TAGTGCATTG AACCCTTTAA 1080

CATTTAAGCA AAATAATAAA CAGTCTTTTG AGGTTCCTTA ACAATGAAAC GTGTTCGAGT 1140

GGCAGCAGCG GAATCCATGC YTCTTCTCCT GGAGTGTGCA AKAGTCCGTG GTCCTGAGTA 1200

TCTCACACAG ATGTGGCATT TTATGTGTGA TGCTCTAATT AAGGCCATTG GTACAGAACC 1260

AGATTCAGAC GTCCTCTCAG AAATAATGCA TTCTTTTGCA AAGGTGAATA TTTTTCTCTT 1320

AAAAAATATG TATAAGGTGG TATGTTCATT TATTAGTCTT GCTAAAAAAA AAAAAAAAAA 1380

ACTTNGAGGG GGGGNCCGGT ACCCAATT 1408

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2031 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

AGGATATGCA TGATTCTTAA CCAGGCTATA TGTTAAAAAA AAATTGGAAA ATGCAATACA 60

TTTTTTACTA TACAAACTAC AGAATGAGTA TGCAAGTTTT ATTTATCAAA ATGTAATGGA 120

TTTTTAAAGG CTGAGAAATT TCCTTATAC CTACCTTTTC AGTTATTTTA ATTATACCAA 180

ATTATCAACT AGAATAGCTT CATCCATATG AAATATAAAA TGAAGAGACA CCTAGGCTCT 240

ATCAGGCTTA GGATTCTTTG AACTTATTTC CACTTTAATT TCTCAGTGGA AGTTAAGAGG 300

GGTGAGAAAA CAAAGAAGGG GAAAAACTGA CAACTAACAA AACCAGCACC ACATCGCTAG 360

GTGGTGCTTA CTAATTACCT TCTCAGGATT TTCCTCAGAT TGAAAAGCTT ATGAGGATTT 420

CTTGGGAGTC TTAATAACCT GCCTGTTAGT ACAGAGCTTT CCTGATGATA TTTACTCTTG 480

AGCACATGTG GTTGTAAAAC CTTAACTTTC TTTCTCCAGG AGGGTGGTGA TAGAAACAGA 540

TGGTAGTATT TATGAACTGA TGTTCTCGTG AAATGTTGAG GGTGGGGAGA AAAGACTTTA 600

AGGGAGGAGA GCCATCTATT TTGTTCCTAA AGCCACCTCT CAGCAGAATC GTCATGTTTT 660

TCTGATGCAC CGCTCTGCTT CATGCCCAAG ATGACTTGCG AGGCAATCTC AGGAGCTGTG 720

GACTTAACCR TTGCAAAGCA CACTGTCTTT CTCAGCGTTC TCTGCAAGTC AGTAGGTGTT 780

AGTATGGTTG CAAAGTTCAC TGTCTCAGCA AAGTTGAACT GGGCTACCTC TCTACAGCTG 840

TTTCCTCAGA GGGAAAAATC TTGAGACCAG ATGGTGGAGC TCTGGAGTCA GAGGAAATGG 900

GTGTCTTCAG CACAAAGCTG CTGCTTTTAC TTCAGCCACT TCTGACATTT TTACATACCG 960

AGCCTGAGAT TRTGTGATTA TCTCAAATCA AATCACTTTG ATGGAGATAA ATAATCAAAA 1020

CTGTTTTATA GTCATTGATT TGGTGAGAAC AGTAATGGAA AATGGTGTTG AAGGACTTCT 1080

CATTTTTGGA GCTTTCCTTC CAGAGTCCTG GCTGATTGGT GTTCGCTGTT CATCTGAGCC 1140

CCCAAAAGCA TTATTACTGA TACTTGCACA CAGTCAAAAG CGCAGACTGG ATGGATGGTC 1200

TTTTATAAGG CATTTAAGGG TACACTACTG TGTTTCACTG ACCATACATT TTTCTTAGCC 1260

CCTCAAGTAA TATAGCACAG AGTTATGAAT GACAATTCCC CTAACCATTC CTCTTCATAT 1320

CTGCCTCTTC CCCTTACCAT CGTAATTCTC CAAACTGGTC ATAAAGGCAC TCTGTGAAGA 1380

TATTGGGGAC TGACATCTTA AGCTCTCACC TGGCTGCAGT AGGAAAGGCC AAACTGACGA 1440

CAAAAAAAAA ATTCTTTATA AAGATGATAT GGTAACATGT ATCTTTGCCC TGGGTCTGGG 1500

TGGGTCCAGT CAGTCTCAGA TTTACAAGCA TTTAGGAGCC TAGGTAAAAG CTGCTAGTAT 1560

TCTTTTAAAA GTTACATTTA TGACTTGCAA TGATAGAAAA CTCCTTCCAA TTAAATGGCA 1620

TTTTATAATA TTATGTGTGT ACTTCACAGT GTTAAAAATA CCCTCATACG TTATTGCATT 1680

TGATCTTCAC AGAAAGTGCA TTTTAACCAG TACTCTGGGT GCAATAAATA ATATGTAGAA 1740

ATTTAAGTCC TCCAATTCCA GCATATCCAG TGAGTTTTGA CAGTGTGTTT ATGTGGAATG 1800

TTTAAGGATA TACAATTGTA CTTTATATAA ATTGGTTCTT GTTCTTCTTA AATGTGACAT 1860

GAAATAATTG TGCTGCTACA TTATACTGGA AATTAACAGG GGAAAAGGGA AGAGCTCTTG 1920

GCTCCCTTGA GGTTCTGCTA GTGGTGTTAG GAGTGGTTAC AACTGAGCTT TTAGTAACCA 1980

TTTAACCGTA TGTAAACTTG GTTTCTAATT AAAAAAAAAT TTCTTTTTCC A 2031

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 971 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

CGCGTCGGAA CTCGGCCGCG GGACATCCAC GGGGCGCGAG TGACACGCGG GAGGGAGAGC 60

AGTGTTCTGC TGGAGCCGAT GCCAAAAACC ATGCATTTCT TATTCAGATT CATTGTTTTC 120

TTTTATCTGT GGGGCCTTTT TACTGCTCAG AGACAAAAGA AAGAGGAGAG CACCGAAGAA 180

GTGAAAATAG AAGTTTTGCA TCGTCCAGAA AACTGCTCTA AGACAAGCAA GAAGGGAGAC 240

CTACTAAATG CCCATTATGA CGGCTACCTG GCTAAAGACG GCTCGAAATT CTACTGCAGC 300

CGGACACAAA ATGAAGGCCA CCCCAAATGG TTTGTTCTTG GTGTTGGGCA AGTCATAAAA 360

GGCCTAGACA TTGCTATGAC AGATATGTGC CCTGGAGAAA AGCGAAAAGT AGTTATACCC 420

CCTTCATTTG CATACGGAAA GGAAGGCTAT GCAGAAGGCA AGATTCCACC GGATGCTACA 480

TTGATTTTTG AGATTGAACT TTATGCTGTG ACCAAAGGAC CACGGAGCAT TGAGACATTT 540

AAACAAATAG ACATGGACAA TGACAGGCAG CTCTCTAAAG CCGAGATAAA CCTCTACTTG 600

CAAAGGGAAT TTGAAAAAGA TGAGAAGCCA CGTGACAAGT CATATCAGGA TGCAGTTTTA 660

GAAGATATTT TTAAGAAGAA TGACCATGAT GGTGATGGCT TCATTTCTCC CAAGGAATAC 720

AATGTATACC AACACGATGA ACTATAGCAT ATTTGTATTT CTACTTTTTT TTTTTAGCTA 780

TTTACTGTAC TTTATGTATA AAACAAAGTC ACTTTTCTCC AAGTTGTATT TGCTATTTTT 840

CCCCTATGAG AAGATATTTT GATCTCCCCA ATACATTGAT TTTGGTATAA TAAATGTGAG 900

GCTGTTTTGC AAACTTAAAA AAAAAWWAAA AAAACTSGAG GGGGGCCCGT ACCCAANTCG 960

CCGNATATGA T 971

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1792 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

GAACCCCCTT TCTCCTGGTA AAGGGTAAGG GGGGGGATAA TGTTTACCAC AGGTACGAAA 60

TAGTCACTTT AACATTGAGA CCTCTGCCTC ATTGAATTCA GGTTTTTTAA GTACTTGAAA 120

CTCTTCAGAT TCTCCTTATT TTAGTTTCTT TTTACATTTA TGAAGTAGAA AGCATTGTTT 180

TGTAAACTGT TTTGAAAATA AATAGCCTAG TCTCTTATCC TCTTTAGCGT GGATTAAAGG 240

TGAAGTTCTG CAAATGGGAG AGTGTTCACA GTAGATAGCT CAGATTGATT GAACACATTT 300

GAGGAAGAGA CTCCTGCATG AGATACCAGC ATTTTTACAA ATACTTTTTA TGTACATTCT 360

TTATTTTGTC ATTTTGTCAA CCCTCTCCCC AAGCACATCT TCTTTCCTTT TACTATGTCT 420

ATGTAGGGAA AAACAAAACA AAAAATTGCA CTTACGTTAC ACTCCCAAAA TGTGGGTAAT 480

CCGTGTCTTT CAAAAAACAT TTCTGTTTTT TGTTTTGTTT TGGTCAGTCC ATTGCATAAG 540

TGACAAGTTT GGGTGCTTGT GGCACGTATG TATGAAGCGG GAGGGGGATG ASAATTGCCT 600

GTCCTTCAGT ARGCTGTAAA AGTAATTTAC ATGTAAGTAA AAAGGGAAAA TAGAATAGAT 660

GCCAAAGTCA TTTATTCAGT CCTTAGTTTT CTTATGTGGC ATTACTGCAT CTGCTAGTTA 720

GTGAGAAAGC ACCCTCAGCT TTTACTGCTC CCCTCCCTGC CTGCCAACAC ACTTGATGTG 780

TGCAAACAGC CCTCAAGTAT CTGTCAGATG ACCTATATAA GGTATTGAAT AAGGTATTCT 840

TGTCAGTTTA GAAATGGACT GGATAAAACT TACTTGGTTG TCATTATTTT ATCTCATTTG 900

TCCTGTTACA TGCCCTATGT TAAGATAATT ATATTGCCAC TAATAATCAA GATGCTAAAT 960

GAGTATTACA ACTGGCTAAT ATCATTTTTT ATATACAAGG GTATGTGTAT ATTTGGAATT 1020

GRTATGAGAA ACTCATTTGT ACCCATTTGA GTGATATTGC ACAACAAACA CAGATAYCTA 1080

CAGACTCCGT TTTCATTTTC TCGTGTTCTT TATGATAATG ATCTTTGTAG ATTGGTTATT 1140

TCTGTACTTT ATCTGTAATA AACTTTGTAG ATCCTGTGAA CCATTACTTT GCCTAAATCA 1200

CTTGAGACTT GAGTCTTTAA TAACAAAGCA TCAATATTCA CTAAAGTCAA TCTCTTTTGA 1260

GTTTCTGTGA CTTGGCTAGA AGCTCTTGAC ACTAAGGGAT TAGTGTTAAT TTTCCCTGGG 1320

GGTGTTCCAC TAGGGCATTA CTGTATAATG ACTTGATGTT GCCACATAGA CTTCAAGATA 1380

TATAATATTT TGAGGATTTT GTTGATTGGC CTATGTTTTA TTGCATAGTG TGAAACGTGT 1440

AAAGCTTGGT TAACCTGTAT ATAGATAGCT TATTGTTGAC TAGTTATAGT GTATTTAGGG 1500

TTGCCTGTAA TATTTAAGCT TCTTTACTGA TGTGTGTGCT GGTAGGAACA TATAATTTTT 1560

GTACATTATA TTTACTGAGA TGTTGCCTTT TTTATTTTAC AAATACTTTG GAATTCCAAT 1620

GTGTTTTTTG CTTCCGTGAG GATTAATTTG GAAAGGTTTT TAATGACATT CCACTGATTT 1680

CAGATTTTGC TTGAGATTGA CTTCAATAAA TTGTCCTGTA TGTTCCAAAA AAAAATTAAA 1740

AAACTCGAGG GGGGCCCGGT ACCCAANNCG CCGGATATGA TCGTAAACAA TC 1792

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 896 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

AGTTGNANAC AACAGGACCT GAGTCCTTGG GCAGCACCAG TAGGTTGCCC CYTGCYTCYT 60

GCCAGCYTCA CYTGCCACYT TYTGCCCCTY TCGGGATGCC TTCGCAGACA GAGYTYTTCG 120

CTGCCTGTGG TGGCCAYTCT TTGCTTTTGG TTYTCTTGCC CCTTGGCCTC CCTTTTTGTC 180

CCCGGGCAGC CTTGTGTGAC CTGCCCTTTT CCCTCCCTTC CTTTCCAGGA CAAGCACGCC 240

GAGGAGGTGC GGAAAAACAA GGAGCTGAAG GAAGAGGCCT CCAGGTAAAG CCTAGAGGCC 300

AAAGAACTTT CCAGGTCAGC CGGACAGCTC CAGCAGCTCC ACGTTCCAGG CAGCCTCGMC 360

CGCCGGCTGC GCTCCCAGCA CTGGGGTTTG GGGGGAGGGG GGTGGCCAAG GGGCGTTTCC 420

TCTGCTTTTG GTGTTTGTAC ATGTTAAGAA TTGACCAGTG AAGCCATCCT ATTTGTTTCC 480

GGGGAACAAT GACGGGGTGG GARAGGGGAG AGGAGAGAGT TTGGGAAAGG GAGATGGAGA 540

AGAACTCAAG GACATTGCAA CCCTGCCCGG CGCAGATCTG ATTTTCACAT CTCTACCTGG 600

ACATTGAGCC TCCCAGGCAC CATGTTGAGG AGAGATGAAA ACCAGGGCGG TAGAACTTCA 660

GGGTGAAGGA CAGAGTCCTG GGTGGGGCAG CGGCTGCAGG GCGCACCAGA GAACCCAGCC 720

AGAGGGGGTG TGAGTACCAG TGGTGTTGCT TCCACCCTGC AGCAGGTGGG ATGAGGTCTG 780

TGTGTGTGTG TGAACCATCA TTTTTTGATC ATCATGACCA ATGAAACATT GAAAAAAAAA 840

AAAAAAACTG GAGGGGGGCC CGTACCCAAN TCGCCGNATA GTGATCGTAA ACAATC 896

(2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 912 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

TCGACCCACG CGTCCGGTCA GCCAGTCGCA TCCAGCCATG ACAGCCTTCT GCTCCCTGCT 60

CCTGCAAGCG CAGAGCCTCC TACCCAGGAC CATGGCAGCC CCCCAGGACA GCCTCAGACC 120

AGGGGAGGAA GACGAAGGGA TGCAGCTGCT ACAGACAAAG GACTCCATGG CCAAGGGAGC 180

TAGGCCCGGG GCCAKCCGCG GCAGGGCTCG CTGGGGTCTG GCCTACACGC TGCTGCACAA 240

CCCAACCCTG CAGGTCTTCC GCAAGACGGC CCTGTTGGGT GCCAATGGTG CCCAGCCCTG 300

ARGGCAGGGA AKGTCAACCC ACCTGCCCAT CTGTGCTGAG GCATGTTCCT GCCTACCATC 360

CTCCTCCCTC CCCGGCTCTC CTCCCAGCAT CACACCAGCC ATGCAGCCAG CAGGTCCTCC 420

GGATCACYGT GGTTKGGTGG AGGTCTGTCT GCACTGGGAG CCTCARGARG GCTCTGCTCC 480

ACCCACTTGG CTATGGGAGA GCCAGCAGGG GTTCTGGAGA AAAAAACTGG TGGGTTAGGG 540

CCTTGGTCCA GGAGCCAGTT GAGCCAGGGC AGCCACATCC AGGCGTCTCC CTACCCTGGC 600

TCTGCCATCA GCCTTGAAGG GCCTCGATGA AGCCTTCTCT GGAACCACTC CAGCCCAGCT 660

CCACCTCAGC CTTGGCCTTC ACGCTGTGGA AGCAGCCAAG GCACTTCCTC ACCCCYTCAG 720

CGCCACGGAC CTYTYTGGGG AGTGGCCGGA AAGCTCCCSG GCCTYTGGCC TGCAGGGCAG 780

CCCAAGTCAT GACTCAGACC AGGTCCCACA CTGAGCTGCC CACACTCGAG AGCCAGATAT 840

TTTTGTAGTT TTTATKCCTT TGGCTATTAT GAAAGAGGTT AGTGTGTTCC CTGCAATAAA 900

CTTGTTCCTG AG 912

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1382 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

AATTCGGCAC GAGCGGAGGC GAGGGAAACT RAGGGCGAAA GTTGTGTGTC GTGTTGGCAG 60

GAGGGCCTAG AAGGGAAAGA CTGTCTAGTG GGACAATGTC ATATTATAAA TTTGGAATGC 120

TGAATAGAAA ATTATAGATT TTGATATTGA AGGAAATGAA GCGAAGCYTA AATGAAAATT 180

CAGCTCGAAG TACAGCAGGC TGTTTGCCTG TTCCGTTGTT CAATCAGAAA AAGAGGAACA 240

GACAGCCATT AACTTCTAAT CCACTTAAAG ATGATTCAGG TATCAGTACC CCTTCTGACA 300

ATTATGATTT TCCTCCTCTA CCTACAGATT GGGCCTGGGA AGCTGTGAAT CCAGAGTTKG 360

CTCCTGTAAT GAAAACAGTG GACACCGGGC AAATACCACA TTCAGTTTCT CGTCCTCTGA 420

GAAGTCAAGA TTCTGTCTTT AACTCTATTC AATCAAATAC TGGAAGAAGC CAGGGTGGTT 480

GGAGCTACAG AGATGGTAAC AAAAATACCA GCTTGAAAAC TTGGRATAAA AATGATTTTA 540

AGCCTCAATG TAAACGAACA AACTTAGTGG CAAATGATGG AAAAAATTCT TGTCCAATGA 600

GTTCGGGAGC TCAACAACAA AAACAATTAA GAACACCTGA ACCTCCTAAC TTATCTCGCA 660

ACAAAGAAAC CGAGCTACTC AGACAAACAC ATTCATCAAA AATATCTGGC TGCACAATGA 720

GAGGGCTAGA CAAAAACAGT GCACTACAGA CACTTAAGCC CAATTTTCAA CAAAATCAAT 780

ATAAGANACA AATGTTGGAT GATATTCCAG AAGACAACAC CCTGAAGGAA ACCTCATTGT 840

ATCAGTTACA GTTTAAGGAA AAAGCTAGTT CTTTAAGAAT TATTTCTGCA GTTATTGAAA 900

GCATGAAGTA TTGGCGTGAA CATGCACAGA AAACTGTACT TCTTTTTGAA GTATTAGCTG 960

TTCTTGATTC AGCTGTTACA CCTGGCCCAT ATTATTCGAA GACTTTTCTT ATGAGGGATG 1020

GGAAAAATAC TCTGCCTTGT GTCTTTTATG AAATCGATCG TGAACTTCCG AGACTGATTA 1080

GAGGCCGAGT TCATAGATGT GTTGGCAACT ATGACCAGAA AAAGAACATT TTCCAATGTG 1140

TTTCTGTCAG ACCGGCGTCT GTTTCTGAGC AAAAAACTTT CCAGGCATTT GTCAAAATTG 1200

CAGATGTTGA GATGCAGTAT TATATTAATG TGATGAATGA AACTTAAGTA GTGATAAAAG 1260

GAAGTTTAGC ATAAATTATA GCAGTTTTCT GTTATTGCTT AATTTACCAT CTCCATAGTT 1320

TTATAGCTAC TATTGTATTT CACTTGTTGA ATTAAAGTAT TTGAATTCTT TTAAAAAAAA 1380

AA 1382

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 872 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

GGGCTACTTC AAAGCCCTGG GCCTTATTTC TTCAGGTAAA AAAATATAAA GTCAGATCTC 60

ATCCCGGCTG GCCATGCTGT TAGACCCTTT CATCCTTCTC TTCTGCCTCT TCTCAACAGC 120

TGCCCAGTCC TGTTTGGAAT TCATATACAT ACAGTTCTAA TACTGATGTA TTTACCCTCA 180

TAAGCCACTC AACCCAGAAT CTTATTTGAA TTATAATCCA GAAACATCAG GTGACGTGTG 240

AGACTACTGT ATGAGAAAGA GACAGTTTAA GGGTCAGTCC AATGGAAAAA AGAGTTCTCA 300

GAGCTTTCTT TAGCTTATTC TCATCAAAGA GCTTTCTCTG CAGAAGGAAC CTACTGGTTC 360

CTCCTTTCCA GTCCTAGAAA TCCTGACCTA GAGTGGCTTA ATCCTGCTAG CACCTCTCTC 420

TCGCACTCTG GTGCCAAATG ACTCCAGGAA CTGGGCCATG ATGTGGTGGG AATGACCTTA 480

CCCTGAGCAT GTCACTCATG CATTGAACAA CAGCTAAGAG CAGAGCTTAG AGCTTAGAGC 540

TGGGCCCTGT AAGGTGAGAG GAATCACATC CTGCAGAAGT CTGTCCTGAG AAGCAGGTAC 600

TCCTGTCACA GCAGAGACAC AGTGGATACC TGAGTAACAA TAATACAAGA CAGGACGTGG 660

GMACAGCAAA AGATTTGGGT GTCAGAAGAR GCCGAGAACA CTTYCAGGCA GGAACATTCA 720

RARTTGTTCT TGGAGGAART AGGCMCSAAG GCTGGGCAGG ATTTCMCGGG GCAGAGATGG 780

AGCAAGCAAT TGAAATGAAA GCCATGGCAT GGGAAAAGGA GCACTGGCCA CAGGGAGTGC 840

AACGTTGTGA TGCAAGGCCA CTGTGGAGCC AT 872

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 812 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

GGCAGAGGCT CACCCCAGCA GAGATTGAGG GGGAACCGTG ATGAAATTTT TAAGTATTCT 60

GCTTGATGAT AATAATTTTY CTCTTATGTT AATGTTGGCT CCGTTTGGGT GTTTAGCTTT 120

TGAAAGGAGT ATGAAAATGC GGAATGGGGC TTTGGGGCTT GAGGAGGTGT GATCTCTAGT 180

GTTTAAAAAA TTTAATTGCA CAAATAGAAA TAATTCACCC ACATTATTGA ACCCCACTAA 240

AGCATATCCT TTTTGTCCAT ATTCCTTTCC TGCTGCCCTC GTGTGTACCA TTATTACTCA 300

GTTGTGATTT GAGCTCGTTC CACTTAAAGT CATTCATAGA TACTTTTGCG TCGTGTTKGA 360

ATATTTATTG AATTTCTATT CTGTGTTTTA CTTAATTACT TTATTATGGA ACCTTTACAC 420

AGGTCTGGTG TACTTGTTCT TTGAAAAGTC TTATGTTGAC CACCATCACT GAGCATATAG 480

CTTTTTCCTT ATTTCCTTGG GATAATTACC CGAAGTGGAA ATACCGAATC AAACTTCTGT 540

TTTCTTTCTT TGGCACTATT ATATAAATTG TTTTCCAAAC AAGGCATGTT TACAATAGAC 600

ATTTTTCAAA ATCTGGGTAT TTGTCCTATT TTGCTCTCTG TATGCAGAAT TCAGCGGGGT 660

GCCAAGTCGT TTTCTGTGTG GGTTGAGAGA CAGGCTGTGC AGCCCACTGT TGCATAGGAC 720

TAACTACTAC AAATCATGCT GAGACCGAGC TATTTTTGCT GCTTAGARGC TTTGCAGCCT 780

TGAGTAAGTT TCGNCATCTG GAAACNTTGN AA 812

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1515 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

AATTCGGCAC GAGGGAAATT CAAGCACTTT TCCTAAAAGA AGGGGGAATG GATGCTGAAA 60

CAACACGTNT CCCACAAAGG GAGCAGACAC TGGGCTTGTG AAGCTGCCCC ATACCTTCCC 120

CACAGAACTG GGGTCCGGCC TCCCTGACAT GCAGATTTCC ACCCAGAAGA CAGAGAAGGA 180

GCCAGTGGTC ATGGAATGGG CTGGGGTCAA AGACTGGGTG CCTGGGAGCT GAGGCAGCCA 240

CCGTTTCAGC CTGGCCAGCC CTCTGGACCC CGAGGTTGGA CCCTACTGTG ACACACCTAC 300

CATGCGGACA CTCTTCAACC TCCTCTGGCT TGCCCTGGCC TGCAGCCCTG TTCACACTAC 360

CCTGTCAAAG TCAGATGCCA AAAAAGCCGC CTCAAAGACG CTGCTGGAGA AGAGTCAGTT 420

TTCAGATAAG CCGGTGCAAG ACCGGGGTTT GGTGGTGACG GACCTCAAAG CTGAGAGTGT 480

GGTTCTTGAG CATCGCAGCT ACTGCTCGGC AAAGGCCCGG GACAGACACT TTGCTGGGGA 540

TGTACTGGGC TATGTCACTC CATGGAACAG CCATGGCTAC GATGTCACCA AGGTCTTTGG 600

GAGCAAGTTC ACACAGATCT CACCCGTCTG GCTGCAGCTG AAGAGACGTG GCCGTGAGAT 660

GTTTGAGGTC ACGGGCCTCC ACGACGTGGA CCAAGGGTGG ATGCGAGCTG TCAGGAAGCA 720

TGCCAAGGGC CTGCACATAG TGCCTCGGCT CCTGTTTGAG GACTGGACTT ACGATGATTT 780

CCGGAACGTC TTAGACAGTG AGGATGAGAT AGAGGAGCTG AGCAAGACCG TGGTCCAGGT 840

GGCAAAGAAC CAGCATTTCG ATGGCTTCGT GGTGGAGGTC TGGAACCAGC TGCTAAGCCA 900

GAAGCGCGTG ACCGACCAGC TGGGCATGTT CACGCACAAG GAGTTTGAGC AGCTGGCCCC 960

CGTGCTGGAT GGTTTCAGCC TCATGACCTA CGACTACTCT ACAGCGCATC AGCCTGGCCC 1020

TAATGCACCC CTGTCCTGGG TTCGAGCCTG CGTCCAGGTC CTGGACCCGA AGTCCAAGTG 1080

GCGAAGCAAA ATCCTCCTGG GGCTCAACTT CTATGGTATG GACTACGCGA CCTCCAAGGA 1140

TGCCCGTGAG CCTGTTGTCG GGGCCAGGTA CATCCAGACA CTGAAGGACC ACAGGCCCCG 1200

GATGGTGTGG GACAGCCAGG YCTCAGAGCA CTTCTTCGAG TACAAGAAGA GCCGCAGTGG 1260

GAGGCACGTC GTCTTCTACC CAACCCTGAA GTCCCTGCAG GTGCGGCTGG AGCTGGCCCG 1320

GGAGCTGGGC GTTGGGGTCT CTATCTGGGA GCTGGGCCAG GGCCTGGACT ACTTCTACGA 1380

CCTGCTCTAG GTGGGCATTG CGGCCTCCGC GGTGGACGTG TTCTTTTCTA AGCCATGGAG 1440

TGAGTGAGCA GGTGTGAAAT ACAGGCCTTC ACTCCGTTAA AAAAAAAAAA AAAAAAAAAA 1500

AAAAAAAAAA AAAAA 1515

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 704 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

AAGATGGTGG CGCCCAGAGC TTCGCTCTAT GCTGCTCCCC TGAGAGAGGC GTTTCCATCA 60

ACCAGTTTTG CAAGGAGTTC AATGAGAGGA CAAAGGACAT CAAGGAAGGC ATTCCTCTGC 120

CTACCAAGAT TTTAGTGAAG CCTGACAGGA CATTTGAAAT TAAGATTGGA CAGCCCACTG 180

TTTCCTACTT CCTGAAGGCA GCAGCTGGGA TTGAAAAGGG GGCCCGGCAA ACAGGGAAAG 240

AGGTGGCAGG CCTGGTGACC TTGAAGCATG TGTATGAGAT TGCCCGCATC AAAGCTCAGG 300

ATGAGGCATT TGCCCTGCAG GATGTACCCC TGTCGTCTGT TGTCCGCTCC ATCATCGGGT 360

CTGCCCGTTC TCTGGGCATT CGCGTGGTGA AGGACCTCAG TTCAGAAGAG CTTGCAGCTT 420

TCCAGAAGGA ACGAGCCATC TTCCTGGCTG CTCAGAAGGA GGCAGATTTG GCTGCCCAAG 480

AAGAAGCTGC CAAGAAGTGA CCCTTGCCCC ACCAACTCCC AGATTTCAAA GGAGGTAGTT 540

GCAAAAGCTG TGCCCAAGGG GAGGAAGGAG GTCACACCAA TATGATGATG GTTTTCATGA 600

CTTTGAATGA TATATTTTTG TACATCTAGC TGTATCGAGG CATCAGGCCT GAATAAACAT 660

CCTTTCTTAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAA 704

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1094 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

GGCAGCTTTC TTACAAACCC ATCCTTCTGA AATGTTGCTT CAAATTCATC CTCTGCTCCC 60

CAGTCCCACT ATTCCACACA TACTGTTACT GTTTCTTTAT CCTACTTTCT CAATTTTGGA 120

ACATAGTTGC AGTTACTGCA TTGAATACCT GTGGGTTTGC CTGTTGTTCT GTCTGTCTCT 180

GTGGTTCTTG TAATANTGGA TCCCAGAGAT AAAATGGACA GTTGTNATGC ACAGTTAATT 240

CAGAAACTAG ACCTTACTTG CTGTGTGAAA TACCAACTAA ATTCTCAGTG AACTCAGCTG 300

ANCTTTATCT CCTTTTGTTT CCCCAATTTA TAATTTCAGT TCAGGCCCAG AAAGATGGAA 360

TCCCAGCTAA GAAATACAAG TTACACCCTG TACTAGCAGC CCATGTGTGC ATGTTCTTTA 420

AGTGCTCTTG CAGCTATGTC ATTTATATTG ATTTCCCTGT ATTATTATAA GCAAAGCAAA 480

TTTGAGGAAA AAAACCCATA ATACCACACC TCATTTTTTT CAAGTAATAG GGTCATAAGT 540

CTCATYCTYC ATATAATATG TTGAGTATGC AGTATATTAT GTGTTAGGCT CTGGANAGGC 600

AGAGGTTAGA TCATGTWACA GATCATATCK GATTAGGCAG ATAAACAGTA TTTTAACCTT 660

TTCCTTATTA TATGTAACTT GCTTTCAGGT TTTTTAATGT TACTATTATG TCTTTAATAT 720

ATTATCTTTA TTTGTACTTT TGTATACAGA GTGATTTTCC TTTTTTAAAA AAAATTGTGT 780

CTTTAGGATG GATTCCAAAG ATGTGGAATC AGTAGGTTTA AGGAATATGG ATATTTTGGC 840

TGGCAAGGTG GCTCACACCT GTAATCCCAG CACTTTGGGA GGCTGAGGTG GGTGGATCAC 900

CTGAAGTCAG GAGTTCGAGA CCAGCCTGAC CAACATGGCG AAACCCTGTT TNTACTAAAG 960

ACACACWWAA AATTRGCCAG TGGTGGTGGC ATGTGCTTGT AGTCCCACTT AGCTACTCGA 1020

GAGGCTGAGG CAGGAGAATC GCTTGAACCC GGGAGGCAGA GGTTGCAGTG AGGCAAGATG 1080

GCACCTCTAC ACTC 1094

(2) INFORMATION FOR SEQ ID NO: 43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1321 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

TGGCTTAGGC CATCACCCTT CCCTTGGCTG GAACTACTGG ACAGACCCTT TTGAGATGTG 60

CCTGTGGTGC TGTGGAGATG TGTGTAGTGG TCTTAGCTCT TTGTTGAGCT TGTGTGTGTG 120

TTGTGTAGTC TTAGCTGTAT GCTGAAATTG GGCGTGTGTT GGAGGGCTTC TTAGCTCTTT 180

GGTGAGATTG TATTTCTATG TGTTTGTATC ASCTGAATGT TGCTGGAAAT AAAACCTTGG 240

TTTGTMAAGG CTCATTTTTG TGGGAAGTAA GTAGGGGAAA AGGTCTTTGA GGGTTCCTAG 300

GCTCCTTTGT ACAACAGGAA AATGCCTCAA AGCCTTGCTT CCCAGCAACC TGGGGCTGGT 360

TCCCAGTGCC TGGTCCTGCC CCTTCCTGGT TCTTATCTCA AGGCAGAGCT TCTGAATTTC 420

AGGCCTTCAT TCCAGAGCCC TCTTGTGGCC AGGCCTTCCT TTGCTGGAGG AAGGTACACA 480

GGGTGAAGCT GATGCTGTAC TTGGGGGATC TCCTTGGCCT GTTCCACCAA GTGAGAGAAG 540

GTACTTACTC TTGTACCTCC TGTTCAGCCA GGTGCATTAA CAGACCTCCC TACAGCTGTA 600

GGAACTACTG TCCCAGAGCT GAGGCAAGGG GATTTCTCAG GTCATTTGGA GAACAAGTGC 660

TTTAGTAGTA GTTCAAAGTA GTAACTGCTA CTGTATTTAG TGGGGTGGAA TTCAGAAGAA 720

ATTTGAAGAC CAGATCATGG GTGGTCTGCA TGTGAATGAA CAGGAATGAG CCGGACAGCC 780

TGGCTGTCAT TGCTTTCTTC CTCCCCATTT GGACCCTTCT CTGCCCTTAC ATTTTTGTTT 840

CTCCATCTAC CACCATCCAC CAGTCTATTT ATTAACTTAG CAAGAGGACA AGTAAAGGGC 900

CCTCTTGGCT TGATTTTGCT TCTTTCTTTC TGTGGAGGAT ATACTAAGTG CGACTTTGCC 960

CTATCCTATT TGGAAATCCC TAACAGAATT GAGTTTTCTA TTAAGGATCC AAAAAGAAAA 1020

ACAAAATGCT AATGAAGCCA TCAGTCAAGG GTCACATGCC AATAAACAAT AAATTTTCCA 1080

GAAGAAATGA AATCCAACTA GACAAATAAA GTAGAGCTTA TGAAATGGTT CAGTAAGGAT 1140

GAGTTTGTTG TTTTTTGTTT TGTTTTGTTT TGKTTTTTTA AAGACGGAGT CTCGCTCTGT 1200

CACTCAGGCT GGAGTGCAGT GGTATGATCT TGGCTCACTG TAACCTCCGC CTCCCGGGTT 1260

CAAGCCATTC TCCTGCCTCA GTCTCCTGAG TAGCTGGGAT TACAGGTGCG TGCCACCATG 1320

CCTGGCTAAT TTTTGTGTTT TTAGTAGAGA CAGGGTTTCA CCATGTTGGT CGGGCTGGTC 1380

TCAAACTCCT GACCTCTTGA TCCGCCTGCC TTGGCCTCCC AAAGTGATGG GATTACAGAT 1440

GTGAGCCACC CGTGCCCTAG CCAAGGATGA GATTTTTAAA GTATGTTTCA GTTCTGTGTC 1500

ATGGTTGGAA GACAGAGTAG GAAGGATATG GAAAAGGTCA TGGGGAAGCA GAGGTGATTC 1560

ATGGCTCTGT GAATTTGAGG TGAATGGTTC CTTATTGTCT AGGCCACTTG TGAAGAATAT 1620

GAGTCAGTTA TTGCCAGCCT TGGAATTTAC TTCTCTAGCT TACAATGGAC CTTTTGAACT 1680

GGAAAACACC TTGTCTGCAT TCACTTTAAA ATGTCAAAAC TAATTTTTAT AATAAATGTT 1740

TATTTTCACA TTGAAAAAAA AAAAAAATTT AAAAACYCGG GGGGGGCCCS GWACCCCATT 1800

NGCCCCTAAG GGGGGGGGTT T 1821

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1024 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

GGGGCACAGT TGAAGAAGCG ACCGAGGGAC TGGGAGTCGT TAGTGAGGAT GACGCGGCAT 60

GGCAAGAACT GCACCGCAGG GCCGTCTACA CCTACCACGA GAAGAAGAAG GACACAGCGG 120

CCTCGGGCTA TGGGACCCAG AACATTCGAC TGAGCCGGGA TGCCGTGAAG GACTTCGACT 180

GCTGTTGTCT CTCCCTGCAG CCTTGCCACG ATCCTGTTGT CACCCCAGAT GGCTACCTGT 240

ATGAGCGTGA GGCCATCCTG GAGTACATTC TGCACCAGAA GAAGGAGATT GCCCGGCAGA 300

TGAAGGCCTA CGAGAAGCAG CGGGGCACCC GGCGCGAGGA GCAGAAGGAG CTTCAGCGGG 360

CGGCCTCGCA GGACCATGTG CGGGGCTTCC TGGAGAAGGA GTCGGCTATC GTGAGCCGGC 420

CCCTCAACCC TTTCACAGCC AAGGCCCTCT CGGGCACCAG CCCAGATGAT GTCCAACCTG 480

GGCCCAGTGT GGGTCCTCCA AGTAAGGACA AGGACAAAGT GCTGCCCAGC TTCTGGATCC 540

CGTCGCTGAC GCCCGAAGCC AAGGCCACCA AGCTGGAGAA GCCGTCCCGC ACGGTGACCT 600

GCCCCATGTC AGGGAAGCCC CTGCGCATGT CGGACCTGAC GCCCGTGCAC TTCACACCGC 660

TAGACAGCTC CGTGGACCGC GTGGGGCTCA TCACCCGCAG CGAGCGCTAC GTGTGTGCCG 720

TGACCCGCGA CAGCCTGAGC AACGCCACCC CCTGCGCTGT GCTGCGGCCC TCTGGGGCTG 780

TGGTCACCCT CGAATGCGTG GAGAAGCTGA TTCGGAAGGA CATGGTGGAC CCTGTGACTG 840

GAGACAAACT CACAGACCGC GACATCATCG TGCTGCAGCG GGGCGGTACC GSTTCGCGGG 900

CTCCGGAGTG AAGCTGCAAG CGGAGAAATC ACGGCCGGTG ATGCAGGCCT GAGTGTGTGC 960

GGGAGACCAA ATAAACCGGC TTGGGTGCGC AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1020

AAAA 1024

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 983 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

CGACACGGCT GCGAGAAGAC GACAGAAGGG CCCGACCGCG AGCCGTCCAG GTCTCAGTGC 60

TGTGCCCCCC CCAGAGCCTA GAGGATGTTT CATGGGATCC CAGCCACGCC GGGCATAGGA 120

GCCCCTGGGA ACAAGCCGGA GCTGTATGAG GAAGTGAAGT TGTACAAGAA CGCCCGGGAG 180

AGGGAGAAGT ACGACAACAT GGCAGAGCTG TTTGCGGTGG TGAAGACAAT GCAAGCCCTG 240

GAGAAGGCCT ACATCAAGGA CTGTGTCTCC CCCAGCGAGT ACACTGCAGC CTGCTCCCGG 300

CTCCTGGTCC AATACAAAGC TGCCTTCAGG CAGGTCCAGG GCTCAGAAAT CAGCTCTATT 360

GACGAATTCT GCCGCAAGTT CCGCCTGGAC TGCCCGCTGG CCATGGAGCG GATCAAGGAG 420

GACCGGCCCA TCACCATCAA GGACGACAAG GGCAACCTCA ACCGCTGCAT CGCAGACGTG 480

GTCTCGCTCT TCATCACGGT CATGGACAAG CTGCGCCTGG AGATCCGCGC CATGGATGAG 540

ATCCAGCCCG ACCTGCGAGA GCTGATGGAG ACCATGCACC GCATGAGCCA CCTCCCACCC 600

GACTTTGAGG GCCGCCAGAC GGTCAGCCAG TGGCTGCAGA CCCTGAGCGG CATGTCGGCG 660

TCAGATGAGC TGGACGACTC ACAGGTGCGT CAGATGCTGT TCGACCTGGA GTCAGCCTAC 720

AACGCCTTCA ACCGCTTCCT GCATGCCTGA GCCCGGGGCA CTAGCCCTTG CACAGAAGGG 780

CAGAGTCTGA GGCGATGGCT CCTGGTCCCC TGTCCGCCAC ACAGGCCGTG GTCATCCACA 840

CAACTCACTG TCTGCAGCTG CCTGTCTGGT GTCTGTCTTT GGTGTCAGAA CTTTTGGGCC 900

GGGCCCCTCC CCACAATAAA GATGCTCTCC GACCTTCAAA AAAAAAAAAA AAAAAAAAGR 960

KGSGGCCGGT CCCCANTCCC CCC 983

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2421 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

CCGGCTGATC GCTGCCGCTC CGCCAATACA ATAGAGCCAK CCACTACCAG CAGCCTGGCC 60

CTCTTCCTCC TTCTCCAGAG AGACCAATCC AGCCGAACTC GGGGTTTGCC TGAGGAGAAG 120

GAGGAAGTGA CCATGGACAC AAGTGAAAAC AGACCTGAAA ATGATGTTCC AGAACCTCCC 180

ATGCCTATTG CAGACCAAGT CAGCAATGAT GACCGCCCGG AGGGCAGTGT TGAAGATGAG 240

GAGAAGAAAG AGAGCTCGCT GCCCAAATCA TTCAAGAGGA AGATCTCCGT TGTCTCAGCT 300

ACCAAGGGGG TGCCAGCTGG AAACAGTGAC ACAGAGGGGG GCCAGCCTGG TCGGAAACGA 360

CGCTGGGGAG CCAGCACAGC CACCACACAG AAGAAACCTT CCATCAGTAT CACCACTGAA 420

TCACTAAAGA GCCTCATCCC CGACATCAAA CCCCTGGCGG GGCAGGAGGC TGTTGTGGAT 480

CTTCATGCTG ATGACTCTCG CATCTCTGAG GATGAGACAG AGCGTAATGG CGATGATGGG 540

ACCCATGACA AGGGGCTGAA AATATGCCGG ACAGTCACTC AGGTAGTACC TGCAGAGGGC 600

CAGGAGAATG GGCAGAGGGA AGAAGAGGAA GAAGAGAAGG AACCTGAAGC AGAACCTCCT 660

GTACCTCCCC AGGTGTCAGT AGAGGTGGCC TTGCCCCCAC CTGCAGAGCA TGAAGTAAAG 720

AAAGTGACTT TAGGAGATAC CTTAACTCGA CGTTCCATTA GCCAGCAGAA GTCCGGAGTT 780

TCCATTACCA TTGATGACCC AGTCCGAACT GCCCAGGTGC CCTCCCCACC CCGGGGCAAG 840

ATTAGCAACA TTGTCCATAT CTCCAATTTG GTCCGTCCTT TCACTTTAGG CCAGCTAAAG 900

GAGTTGTTGG GGCGCACAGG AACCTTGGTG GAAGAGGCCT TCTGGATTGA CAAGATCAAA 960

TCTCATTGCT TTGTAACGTA CTCAACAGTA GAGGAAGCTG TTGCCACCCG CACAGCTCTG 1020

CACGGGGTCA AATGGCCCCA GTCCAATCCC AAATTCCTTT GTGCTGACTA TGCCGAGCAA 1080

GATGAGCTGG ATTATCACCG AGGCCTCTTG GTGGACCGTC CCTCTGAAAC TAAGACAGAG 1140

GAGCAGGGAA TACCACGGCC CCTGCACCCC CCACCCCCAC CCCCGGTCCA GCCACCACAG 1200

CACCCCCGGG CAGAGCAGCG GGAGCAGGAA CGGGCAGTGC GGGAACAGTG GGCAGAACGG 1260

GAACGGGAAA TGGAGCGGCG GGAGCGGACT CGATCAGAGC GTGAATGGGA TCGGGACAAA 1320

GTTCGAGAAG GGCCCCGTTC CCGATCAAGG TCCCGTRACC GCCGCCGCAA GGAACGTGCG 1380

AAGTCTAAAG AAAAGAAGAG TGAGAAGAAA GAGAAAGCCC AGGAGGAACC ACCTGCCAAG 1440

CTGCTGGATG ACCTTTTCCG AAAGACCAAG GCAGCTCCCT GCATCTATTG GCTCCCACTG 1500

ACTGACAGCC AGATCGTTCA GAAAGAGGCA GAGCGGGCCG AACGGGCCAA GGAGCGGGAG 1560

AAGCGGCGAA AGGAGCAAGA AGAAGAAGAG CAAAAGGAGC GGGAGAAGGA AGCCGAGCGG 1620

GAACGGAACC GACAGCTGGA GCGAGAGAAA CGTCGGGAGC ACAGTCGGGA GAGGGACAGG 1680

GAGAGAGAGA GAGAAAGGGA GCGGGACAGG GGGGACCGAG ATCGGGATAG GGAAAGGGAC 1740

CGAGAACGAG GCAGGGAAAG GGATCGCAGG GACACCAAGC GCCACAGCAG AAGCCGGAGT 1800

CGGAGCACAC CTGTGCGGGA CCGGGGTGGG CGCCGCTAGC TGGGAAAACA CTAGAGCTGC 1860

AGGTACCAGC CACTCGGCCC CAGGGGGTTA TGGCCACAGA GGGATAGGCA CAGTCTCCAC 1920

CACCCTGGAG CCAAGGGTCT TTCACATCAC CTATCCCTAC ATACATACCA AATGGAAAAG 1980

TGGCCATCCT TTTCCCCCCA AACACACCCC CTTAACCTAT CTCTTGGGAC TTAGCCCGAC 2040

CCTCCCTCTC ATTTCCCATT AAGTCTGAGA GGCAAGAGCT AGGTTAGGCA AGGAGGTGGT 2100

TGGCCAGAGA TGGGGAACAG CCAGGTGCCC CAGTCCTCTG ATTTTTCCTC CATCCTGCTT 2160

ACCACCTCCC TGGGTACTTA CAGCCTTCTC TTGGGAACAG CCGGGGCCAG GACTGGGTCA 2220

CCTATGAGCT GAATCAGCAT CTCCTCCTGA GTCCCAGGGC CCCTGCAGTT CCCAGTCTCT 2280

TCTGTCCTGC AGCCCTTGCC TCTTTCCCAC AGGTTCCACT TTATATCCAC CTTTTCCTTT 2340

TGTTCAATTT TTATTTTTAT TTTTTTTATT ATTAAATGAT GTGGTCTATG GAAAAAAAAA 2400

TAAAAATCTG ACTTAGTTTT A 2421

(2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 840 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

CTCAAACTCC TGAGCTGAAG CGATCTACCT GCCTCAGCTA GGATTACAGG TGTGAGCCAC 60

CGCACCCAAC CTCAATAAGC KTATTTGATA AAAKATATGC AAGCTCCCTT TATKCACTTT 120

TCATTCAGAA TGTTTAGTAA TTTGTATTGT TTTTCAGATT TTCAGCCCAA TATATCTCCY 180

TGCCCACTGT GTCACTGTAT TCTACCTAWA CATCATCACG TGTTTCTGCT ATTGGCTGTA 240

TGATGGAACA CTGCGGCTCA TTTTCCTGAA AACTGCCGAT AGTGCATAGA RTGCTGGGAT 300

GGAAACCAGA ARCTTTGAAT TCAAGCCTTG GTTCTGCCTT GTTTTTGCTT GGGTGGCCTT 360

GAGTCAGCCA CATACCTTTT AAAATCTCAA TTTATTAGAA ATTATTCCAA ATCAAAATCA 420

AATGAGAAGG TATATACAAA AGTGCTTTAT CCCACAATAA ACTATTCAAG AGAGAGCAAA 480

GGAGAGGACA TTTACTCAAC ACCTCCTAAA AGGCAGCCAG TGAAATTAGG CATTTTATTT 540

AATCCTCCTG GCAACTCTGA GAGTAAAGCA TTATTAATCC CATTTTGGCT GTTTAAAGAA 600

ATTATTTGCA CTAGATTCCA GCTGTAGTTT AGYTTCAGAA AAAAAAATCC TGAGATGTGA 660

ATTCACAGCT TTCTGGGTTT AAAGCCCAAG CTCTATCACA TCATGCTATT ATTGTTACAT 720

TACTGCTAGT TCTATGAAAA GAAATACTAA TTTATGAAAT ACATCTTATC CAAAAAAAAA 780

AAAAAAAAAC TGGGAGGGGG GGCCCGTACC CAAATCGCCG GATAGTGATC GTAAACAATC 840

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2432 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

GGCACGAGGC CCGGAACGCT GAGGAAGGGC CCGTCCCGCC TTCCCCGGCG CGCCATGGAG 60

CCCCGGGCGG TTGCAGAAGC CGTGGAGACG GGTGAGGAGG ATGTGATTAT GGAAGCTCTG 120

CGGTCATACA ACCAGGAGCA CTCCCAGAGC TTCACGTTTG ATGATGCCCA ACAGGAGGAC 180

CGGAAGAGAC TGGCGGASTG CTGGTCTCCG TCCTGGAACA GGGCTTGCCA CCCTCCCACC 240

GTGTCATCTG GCTGCAGAGT GTCCGAATCC TGTCCCGGGA CCGCAACTGC CTGGACCCGT 300

TCACCAGCCG CCAGAGCCTG CAGGCAYTAG CCTGYTATGY TGACATCTCT GTCTCTGAGG 360

GGTCCGTCCC AGAGTCCGCA GACATGGATG TTGTACTGGA GTCCCTCAAG TGCCTGTGCA 420

ACCTCGTGCT CAGCAGCCCT GTGGCACAGA TGCTGGCAGC AGAGGCCCGC CTAGTGGTGA 480

AGCTCACAGA GCGTGTGGGG CTGTACCGTG AGAGGAGCTT CCCCCACGAT GTCCAGTTCT 540

TTGACTTGCG GCTCCTCTTC CTGCTAACGG CACTCCGCAC CGATGTGCGC CANAGCTGTT 600

TCAGGAGCTG AAAGGAGTGC GCCTGCTAAC TGACACACTG GAGCTGACGC TGGGGGTGAC 660

TCCTGAAGGG AACCCCCCAC CCACGCTCCT TCCTTCCCAA GAGACTGAGC GGGCCATGGA 720

GATCCTCAAA GTGCTCTTCA ACATCACCCT GGACTCCATC AAGGGGGAGG TGGACGAGGA 780

AGACGCTGCC CTTTACCGAC ACCTGGGGAC CCTTCTCCGG CACTGTGTGA TGATCGCTAC 840

TGCTGGAGAC CGCACAGAGG AGTTCCACGG CCACGCAGTA ASCCTCCTGG GGAACTTGCC 900

CCTCAAGTGT CTGGATGTTC TCCTCACCCT GGAGCCACAT GGAGACTCCA CGGAGTTCAT 960

GGGAGTGAAT ATGGATGTGA TTCGTGCCCT CCTCATCTTC CTAGAGAAGC GTTTGCACAA 1020

GACACACAGG CTGAAGGAGA GTGTAGCTCC CGTGCTGAGC GTGCTGACTG AATGTGCCCG 1080

GATGCACCGC CCAGCCAGGA AGTTCCTGAA GGCCCAGGTG CTGCCCCCTC TGCGGGATGT 1140

GAGGACACGG CCTGAGGTTG GGGAGATGCT GCGGAACAAG CTTGTCCGCC TCATGACACA 1200

CCTGGACACA GATGTGAAGA GGGTGGCTGC CGAGTTCTTG TTTGTCCTGT GCTCTGAGAG 1260

TGTGCCCCGA TTCATCAAGT ACACAGGCTA TGGGAATGCT GCTGGCCTTC TGGCTGCCAG 1320

GGGCCTCATG GCAGGAGGCG GCCCGAGGGC AGTACTCAGA GGATGAGGAC ACAGACACAG 1380

ATGAGTACAA GGAAGCCAAA GCCAGCATAA ACCCTGTGAC CGGGAGGGTG GAGGAGAAGC 1440

CGCCTAACCC TATGGAGGGC ATGACAGAGG AGCAGAAGGA GCACGAGGCC ATGAAGCTGG 1500

TGACCATGTT TGACAAGCTC TCCAGGAACA GAGTCATCCA GCCAATGGGG ATGAGTCCCC 1560

GGGGTCATCT TACGTCCCTG CAGGATGCCA TGTGCGAGAC TATGGAGCAG CAGCTCTCCT 1620

CGGACCCTGA CTCGGACCCT GACTGAGGAT GGCAGCTCTT CTGCTCCCCC ATCAGGACTG 1680

GTGCTGCTTC CAGAGACTTC CTTGGGGTTG CAACCTGGGG AAGCCACATC CCACTGGATC 1740

CACACCCGCC CCCACTTCTC CATCTTAGAA ACCCCTTCTC TTGACTCCCG TTCTGTTCAT 1800

GATTTGCCTC TGGTCCAGTT TCTCATCTCT GGACTGCAAC GGTCTTCTTG TGCTAGAACT 1860

CAGGCTCAGC CTCGAATTCC ACAGACGAAG TACTTTCTTT TGTCTGCGCC AAGAGGAATG 1920

TGTTCAGAAG CTGCTGCCTG AGGGCAGGGC CTACCTGGGC ACACAGAAGA GCATATGGGA 1980

GGGCAGGGGT TTGGGTGTGG GTGCACACAA AGCAAGCACC ATCTGGGATT GGCACACTGG 2040

CAGAGCMANT GTKTTGGGGT ATGTGCTGCA CTTCCCAGGG AGAAAACCTG TCAGAACTTT 2100

CCATACGAGT ATATCAGAAC ACACCCTTCC AAGGTATGTA TGCTCTGTTG TTCCTGTCCT 2160

GTCTTCACTG AGCGCAGGGC TGGAGGCCTC TTAGACATTC TCCTTGGTCC TCGTTCAGCT 2220

GCCCACTGTA GTATCCACAG TGCCCGAGTT CTCGCTGGTT TTGGCAATTA AACCTCCTTC 2280

CTACTGGTTT AGACTACACT TACAACAAGG AAAATGCCCC TCGTGTGACC ATAGATTGAG 2340

ATTTATACCA CATACCACAC ATAGCCACAG AAACATCATC TTGAAATAAA GAAGAGTTTT 2400

GGACAAAAAA AAAAAAAAAA AAAAAAAAAA AA 2432

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1742 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

GTCCTGCAGG AGCTGCACGC GGCCGAGGTG CGCANGAACA AGGAGCAGCG AGAAGAGATG 60

TCGGGCTAAG GGCCCGGSAC GRGSGGCGCC CATCCTGCGA CGGAACACGT TCGGGTTTTG 120

GTTTTGTTTC GTTCACCTCT GTCTAGATGC AACTTTTGTT CCTCCTCCCC CACCCCAGCC 180

CCCAGCTTCA TGCTTCTCTT CCGCACTCAG CCGCCCTGCC CTGTCCTCGT GGTGAGTCGC 240

TGACCACGGC TTCCCCTGCA GGAGCCGCCG GGCGTGRAGA CGCGGTCCCT CGGTGCAGAC 300

ACCAGGCCGG GCGCGGCTGG GTCCCCCGGG GGCCCTGTGA GAGAGGTGGY GGTGACCGTG 360

GTAAACCCAG GGCGGTGGCG TGGGATCRCG GGTCCTTACG CTGGGCTGTC TGGTCAGCAC 420

GTGCAGGTCA GGGCAGGTCC TCTGAGCCGG CGCCCCTGGC CAGCAGGCGA GGCTACAGTA 480

CCTGCTGTCT TTCCAGGGGG AAGGGGCTCC CCATGAGGRA GGGGCGACGG GGGAGGGGGG 540

TGATGGTGCC TGGGAAGCCT GCKTGTGCAN CCGGTGCTTG TTGAACTGGC AGGCGGGTGG 600

GTGGGGGCTG CAGCTTTCCT TAATGTGGTT GCACAGGGGT CCTCTRAGAC CACCTGGCGT 660

GAGGTGGACA CCCTGGGCCT TCCTGGAAGC CTGCAGTTGG GGGCCTGCCC TGAGTCTGCT 720

GGGGAGTGGG CATTCTCTGC CAGGGACCCA TGAGCAGGCT GCATGGTCTA GAGGTTGTGG 780

GCAGCATGGA CAGTCCCCCA CTCAGAAGTG CAAGAGTTCC AAAGAGCCTC TGGCCCAGGC 840

CCCTCCGTGG GACAGCCCCG CCGCCCCTCC CCACCAGGGC TTTGCAGATG TCCTTGAAAG 900

ACCCACCCTA GAGCCCTTTG GAGTGCTGGC CCCTCCTGTG CCCTCTGCCC TGGTGGAAGC 960

GGCASCACAA GTCCTCCTCA GGGAGCCCCA AGGGGGATTT TKTGGGACCG CTGCCCACAG 1020

ATCCAGGTGT TGGAAGGGCA GCGGGTAAGG TTCCCAAGCC AGCCCCAACA CCCTTCCCAC 1080

TTGGCACCCA GAGGGGGCTG TGGGTGGAGG CCTGACTCCA GGCCTCTCCT GCCCACACCC 1140

TCTGGGCTGA GTTCCTTCTT TCCCTTGGAC GCCCAGTGCT GGCCTTGGAG GACGGTCAGC 1200

TGGAGGATGG CGGTGGGGGA GGCTGTCTTT GTACCACTGC AGCATCCCCC ACTTCTCCAC 1260

GGAAGCCCCA TCCCAAAGCT GCTGCCTGGC CCCTTGCTGT AAAGTGTGAA GGGGGCGGCT 1320

GAGTTCTCTT AGGACCCAGA GCCAGGGCCC TCAACTTCCA TCCTGCGGGA GGCCTTGGCC 1380

GGGCACTGCC AGTGTCTTCC AGAGCCACAC CCAGGGACCA CGGGAGGATC CTGACCCCTG 1440

CAGGGCTCAG GGGTCAGCAG GGACCCACTG CCCCATCTCC CTCTCCCCAC CAAGACAGCC 1500

CCAGAAGGAG CAGCCAGCTG GGATGGGAAC CCAAGGCTGT CCACATCTGG CTTTTGTGGG 1560

ACTCAGAAAG GGAAGCAGAA CTGAGGGCTG GGATATTCCT CATGGTGGCA GCGCTCATAG 1620

CGAAAGCCTA CTGTAATATG CACCCATCTC ATCCACGTAG TAAAGTGAAC TTAAAAATTC 1680

AATCAAATGA ACAATTAAAT AAACACCTGT GTGTTTAAGA AAAAAAAAAA AAAAAAACTG 1740

CG 1742

(2) INFORMATION FOR SEQ ID NO: 50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1487 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

GGCACGAGCC TCCGCGAACT GTGGAGTCGG CGGAGGGCTG GAATCAGCGT GGGCTCCAGG 60

TCGCTGGCAG CCGGGTGGCA GAACTCTTCC GAGGCTCCTT GGGAAGAAGC TACACCCGAG 120

GGAGCCGGAT GGGCCTCGAA AACCTGGCCC GCTCTGGTTC TGTACCATTG CAAGGGGAAC 180

CGTAAACTGA GCTTTTCTAA CGTGGGTTTC TGCCAAGTAC TTTTCCAGCT GCCCCCTTCC 240

CCCCAGCACA CAGGAGAGCC TCTGTGTAGC CAGCGCTTGA CAGTCGTTAG GTAGGTTGTA 300

CTGTGTAGGG AGGAGCTCAA GATCATGAAT GGTTGTCACA GGAGAAAGCG GTTGCATCTT 360

TGCAAAACTA TATACCTGCT GTGGTTTGTG TTTTCTTTTC TGCTGAGTAA TGAAGTTGTA 420

AGTTCACACT GGCACATTCT CAGGGCTGTG CAGATTATTT GCACTTTATT TCATAGGTGR 480

ATAAGTGCTT TTTAGCTTTC TTTGTATATT GAGTTGCTTT TGAATTGCTT CCCATATTTT 540

TATTTCATAC AAACTGAACA ATTGTGGCCC CTCTATTTTA TTTATAAAGG TTCAGTGTAT 600

CTTTGCCTGC CTACATCAAT CTGCAAGGGA GTTGCAGAAA GCCTCATGTT CATCGAGCCG 660

TGAGTCACAA CCAATTTCTA AGCTGTTATA ACAAAAAAGT GTTTGCTTTT TTTCACAAGT 720

AACTTTAAAA GTGTAGTTTA GAAAGAAAAC ATTTTCAATA AAAAGACACT ACATTAATCC 780

TGGATGCTTG CAAATCCTAA AATMTATTCC TCCTCTAGCG TTGCACAGCT CTGTGTTGTA 840

TACACAGACT AGCTTTAAAA TTTGTCACAT ACCACTTTAC CTTTACTTTT ATGTATCATT 900

CCCCCGACTT CCTTACTGCA GGTGTGGGCA AGAAAACTTT TCCTTTAACA CTTTTCAACA 960

GCGGGCATAA AATTCTGCAG CTGAGGTCTT GAAGAATGCA GATGGGTACA GTATGTGTTG 1020

GAGCTCACAG TGTGTATTGA CTAACCTAGT TCCTTTTTTG CTTTTTTTGG TATTGTCTTG 1080

TTAAAAGTGA CTCCCAGGTA GCAACTCTCT TTTTTAAGGG TGGGAACGAA AGGGACGTAG 1140

GAAGAATAGA TCTAGATTAT TTAACAGTCT TCGATAGAGT TTGAAAGCTT TCTTCTTCAT 1200

TCAATTTTGG GCAAAATACT GCCTCTGCAT TTGTTCATAA CAAAAAGATT AGATTAATAA 1260

GTAGCTTTTG TTGGTGGAAA TTACCAGCTC TATAAGTCAC CCTTGGTGGT TCATGGACCT 1320

CTGATTAGCT TGGGTTTTGC AGTCTCATTG CCACATGTAT ATGTGGAGCC AATGGCCTTT 1380

TGGTGCTCAG CTGTTTACGT CTGACTCCTT GACTTCTTTG GTACAGTGAT GGAGTCAGAT 1440

CTCATTAAGT GTGATTCTCC ATGGATATAA CCAGCCCCAA AAAAANG 1487

(2) INFORMATION FOR SEQ ID NO: 51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1328 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

GGCACGAGCT CGTGCCGAAT TCGGCACGAG AGAAGATTTG AAGAAGCCAG ATCCAGCTTC 60

CCTGCGGGCT GCTTCTTGTG GGGAAGGGAA AAAGAGGAAG GCCTGTAAGA ACTGCACCTG 120

TGGCCTTGCC GAAGAACTGG AAAAAGAGAA GTCAAGGGAA CAGATGAGCT CCCAACCCAA 180

GTCAGCTTGT GGAAACTGCT ACCTGGGCGA TGCCTTCCGC TGTGCCAGCT GCCCCTACCT 240

TGGGATGCCA GCCTTCAAAC CTGGGGAAAA GGTGCTTCTG AGTGATAGCA ATCTTCATGA 300

TGCCTAGGAG GTTCCTGACA TGGGACCCAT CTGCTCCTCC AGCCAACTCC TGTCCCTCAC 360

ATCCCACCAT GGTGGCTCCT CCCACCTCCT CTGGATTTGT TCACTCTGAG ATCTGTTTGC 420

AGAGTGGGTG CTTAGCAGAC AGAGTGAAGC TGGCTGGGGG GCACAGTGGT GTGTAGTGCT 480

GCTGTGTATC AAAAGACCAA GGTATTATGG GACCTGGTTT CAGAATGGGA TGGGTTTCTT 540

CACCTCATGT TAAGAGAAGG GAGTGTGTCC TGAAGAAGCC CTTCTTCTGA TGTTAAAATG 600

CTGACCAGAA CGCTCTTGAG CCCAGGCATC GTTGAGCATT AACACTCTGT GACAGAGCTG 660

CAGACCCCTG CCTTGAGTCT CATCTCAGCA ATGCTGCCAC CCTCTTGTCT TTCAGAGTTG 720

TTAGTTTACT CCATTCTTTG TGACACGAGT CAAGTGGCTC ACAACCTCCT CAGGGCACCA 780

GAGGACTCAC TCACTGGTTG CTGTGATGAT ATCCAGTGTC CCTCTGCCCC CTTCCATCCC 840

CAACCACATT TGACTGTAGC ATTGCATCTG TGTCCTGTTG TCATTTATGT TAACCTTCAG 900

GTATTAAACT TGCTGCATAT CTTGACATAT CTTGAGATTC TGCATGTCTT GTAAAGAGAG 960

GGGATGTGCA TTTGTGTGTG ATGTTGGATA GTCATCCACG CTCAGTTTGG ACCATTGGAG 1020

GAACTTAGTG TCACGCACAA ATGGGGCTAT TCCTACGCTT AGAATAGGGC TTGTCTGCCC 1080

ACTTTAGAAG AGTCCCAGGT TGGTGAGCAT TTAGAGGGAA GCAGGGCAGA ACTCTGAACG 1140

ACAATACGTC TCTCTGAGCA GAGACCCCTT TGTTCTTGTT ATCCACCCAT ATGGACTTGG 1200

AATCAATCTT GCCAAATATT TGGAGAGATT GTGTGGATTT AAGAGACCTG GATTTTTATA 1260

TTTTACCAGT AAATAAAAGT TTTCATTGAT ATCTGTCCTT GAAAAAAAAA AAAAAAAAAA 1320

AAACTCGA 1328

(2) INFORMATION FOR SEQ ID NO: 52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1856 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

GAATTCGGCA CGAGCTCTGC AACATTGCAA ATGAACTTGC AGTCGAGGGT TCCGCTGCCC 60

CCTAGATTAA ATTCCCCGGG CTGAAACTGA GTTGCAGATT TACAATATCA TATTTTAAAT 120

TGCTGTCTTC AATTAAACCA TTTATGACCA TAACTAATTT TCAGGATGTC GATGCATGCT 180

TTTCCAGGCC TTCCTTCTTT GTACAAAAGT AAATGTCCAT AAAGCGTTTC ACTTATATTC 240

TTCAAACATG ATGCTAATTT AAATTAATTA CTTCCTATGA TATGTTATTA TTCCTATGAT 300

TTTGCCACTG TTATTAGTTC TCTCAAAAAT ACATCTAGGG AAGAGGATTA TTTTAAGTRA 360

TTTGATTATC TTTCTATCTC TTTTATTTAT TTCTCATTTA CTTAAGAAAT TCGTTCCATT 420

GGTTGGCATT GATACAGTAA ATTTGTAAAT GAGGAGACAA TATAAAAAAT CTAAATTACT 480

TGTGCTTAAT GACTGTAGCA GAATSCCTTT TCTCTAAATC AGATTGTCTT TCTTGCAGTT 540

TAGTTTGATA GATTTGCAAG CTATGCTGCT TCCATGAAGT TAGCTGCGCT GGTAGGAACG 600

CAGGCTTCTT TGTCTCTGGT TGTAGCTTGC ATGATCGCCT CATTAGGCAG ACAACGTAGC 660

CGGAGATCAC AAATCAGGCC CTTGGTGTAG TTGCTAGTGT GTGGAGGTGC AGAGAGGTTG 720

GCAGAAACTG ACCTCACTGG GCAAGGGTGG CCATGGACCT GATTCTTTAA TGCACTCTAT 780

GTGTTCAGGA AGCCACAGGC CATATTTGAC TCTGAGAAAG AAAACAAGAG GAAAAACCCC 840

ACAAAGTATA ACAACCCCTT AAGATACATC TATTTTAAAG TGAAATTAAT TTTTCAGTTT 900

ATACCATTGG CCAATTACAA GATAAAAATG TTCAATTTCT TTAAGAATCC TTTGTTGACT 960

TGTCTTTTCA TCTCTTGCTA TTTATATTTG TCACTGTTAG TGAACAAAGT CTTATTTGCT 1020

GAGGAAGGAC TTTGCTGCAC TTACTGTACC ACATCAAACA CTGGGGAGGG TGGTGTTTAA 1080

CTTTTTAAAA AATGTTATTC TGATTATAAC AATAATATTG GCTTTTTTCA TGAAAAGAGC 1140

GCCACCTTGC AAGGTTTAGT GAGATTTATG GAAGTTGAAC ACCTAAGCAG GAATTGCTGC 1200

TAGCTCCAAA AATTTGCGAA GCAAAAGCTA GCCCCAATTG GTTTGGAAGT TTGAAACTGA 1260

TTAACAGATT TGCATTTGAA GTGACTCCAG ACATTAGGTC CAGACATTAG TTAAAAATAG 1320

AAAGAGGAAT AAAGACATCT YTTCTCTCTA GAAAAGATAA CACCTCAACT AATAATCCTT 1380

CCCACTTTCA TTGAGATCAG CTTGTCTGAT AACCTGATAC GAGTGTGACA ATGATAAACA 1440

TGATAATAGT GGTACTTTTG TAATTTTGCT GGTGCATTCA AGAAGATAGT AAAKGATGAG 1500

TTCAYCTTTT CTYCGAACAT YCCTATTCCT AGATGTAGTT CACCTCAAAT TGGGAATTAT 1560

AACTGTCCTA ATTTTTGTTG TGTACCCTGA TGCCCCTTTT GCTTTAATAC CCACAGTGTA 1620

ACAATTAAAT ATCACACTAT GACATACTAT TTAAGTAGGA CATTTTAAAG ATAAATTTTA 1680

GGGGTAAATG TTTACTTCAA AATGACTCCA TATTTCAAAT ATCTGTTTAG ACTGTGAAGG 1740

CCAAATAATT TTTAAGAAAA CATTTGAAGA GTAGTGTGTT TGCATTTGTG AATAATCTTA 1800

CTCACAGCAA GTAAACGTAA TAAAAGCCAA CATTTAAGCC AAAAAAAAAA AAAAAA 1856

(2) INFORMATION FOR SEQ ID NO: 53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1558 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

TGGGTATCCA TTCCTGNAAT TACTTTACTT AGGATAATGG CCTCCAGCTC CGTCCAAGTT 60

GCTGCAAAAG GTATTATTTC GTTCCTTTTT GTGGCTGAGT AGTATTCCAT GGTGTATATA 120

TACCACATTT TCTTTATCCA CTCATTGCTT GATGGGCAGT TAGGTTGGTT CCACATCTTT 180

GCAATTGTGA GTTGTGCTGC TCCAGATATC ATCTTTAACT CCTTTGCCTT CTCCACATAC 240

ATTTCCAAGT CCTGTTCATT CTACCTCCAA AATGTATCTT GTATCCATTC ATCTCTCTCC 300

ATCTTCAATC TATTTCAATG CCCCATCATC TCTTGCATGG AGGAGTGTAA TAATTGGCTA 360

ACTGGCCTGT TCTTACATTT TAAAATCAAA AGATGTGACA GGTGAAATGC CTATTTCAGT 420

GTCCATTGAT GGTTCTGCTT ACACACCACC TGGCTGCCTG GTGTCGCAGT GGCAGAGTTG 480

AGCAGTGTGA AAAAGACTGC TTGGCCCTTT ACAGGGAAAG CAGGTCCACT GTGGCCTGTG 540

AGGACGAGAG CTCTGGGCAG GCTCGGACAC TGGCAGACCC TGGTCCTGGC TGGCCAAGGC 600

AGCAGGGTAT GTGTTTCGGG TCACTCACAG GGCTCAGCAC CACTCCTCAT GGCTTCCTTA 660

CTGTTTCGGC AGAGGCTGAC CCGCGGCTGA TTGAGTCCCT CTCCCAGATG CTGTCCATGG 720

GCTTCTCTGA TGAAGGCGGC TGGCTCACCA GGCTCCTGCA GACCAAGAAC TATGACATCG 780

GAGCGGCTCT GGACACCATC CAGTATTCAA AGCATCCCCC GCCGTTGTGA CCACTTTTGC 840

CCACCTCTTC TGCGTGCCCC TCTTCTGTCT CATAGTTGTG TTAAGCTTGC GTAGAATTGC 900

AGGTCTCTGT ACGGGCCAGT TTCTCTGCCT TCTTCCAGGA TCAGGGGTTA GGGTGCAAGA 960

AGCCATTTAG GGCAGCAAAA CAAGTGACAT GAAGGGAGGG TCCCTGTGTG TGTGTGTGCT 1020

GATGTTTCCT GGGTGCCCTG GCTCCTTGCA GCAGGGCTGG GCCTGCGAGA CCCAAGGCTC 1080

ACTGCAGCGC GCTCCTGACC CCTCCCTGCA GGGGCTACGT TAGCAGCCCA GCACATAGCT 1140

TGCCTAATGG CTTTCACTTT CTCTTTTGTT TTAAATGACT CATAGGTCCC TGACATTTAG 1200

TTGATTATTT TCTGCTACAG ACCTGGTACA CTCTGATTTT AGATAAAGTA AGCCTAGGTG 1260

TTGTCAGCAG GCAGGCTGGG GAGGCCAGTG TTGTGGGCTT CCTGCTGGGA CTGAGAAGGC 1320

TCACGAAGGG CATCCGCAAT GTTGGTTTCA CTGAGAGCTG CCTCCTGGTC TCTTCACCAC 1380

TGTAGTTCTC TCATTTCCAA ACCATCAGCT GCTTTTAAAA TAAGATCTCT TTGTAGCCAT 1440

CCTGTTAAAT TTGTAAACAA TCTAATTAAA TGGCATCAGC ACTTTAACCA AAAAAAAAAA 1500

AAAAAAAAAA AAANAAAAAA AAAAGGGGGC CGCTCTAGAG GTCCAAGTTA NGACGNGG 1558

(2) INFORMATION FOR SEQ ID NO: 54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 948 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

TAAAAATCAT GCTCTGTACC ATCCTCACCG TAGTCATCAT CATCGCCGCG CAGACCACGA 60

GAACTACTGG GATCCCTAAA AACGCCCCTG GTCCGGCCCC ACTCTGCGCC CCTCGATCTC 120

CCAGGCTCTT TCTGCAGWCA TACCGCGGAC CCAATGGGCG CCCTGCACAC CCGTTTCTGG 180

GGCCGTCAGA CTTGGATACA TCGTAAACTC CGCCTCCACG GAACGTCTCG CCTKGCGAGC 240

AAGMTCGGAA TCCAGTTCCT CAGGAACCCC TCCAAAACCC ACACCCCCAG GGACGCCGCT 300

TTCCGGGATC CCGGSCAAAC GCCGGACCCT CAGTCGCTCC AGGCCCCCTC ACCCTCAAAG 360

TGTAGCGCCC CCAACCGAGC AACCTCGGTT TGGTCCCTAA AACCCCGCCT CCTCTATAAG 420

CACCGCCCCA GCTCTGACAA AACCCCGCCT CCAGGTCGGC AGGCTCCGCT TCTTTTCTTC 480

TCCGCGGGGT GATTCAGTCC AGTGATTGGG TTTGTGGCTC CAGGCCTCGC CCACAGACGG 540

ACAGACCCCT CCCTTTCTTC CGGCAAAAGG ACCGAGCCCT GGGGTAGTAA GGSCCCCACA 600

CTCCTGTTTT TTGCAAGTAC ATTTTTGTCC YTCCTCCACC CAGGTATCTG CCTATTTTCT 660

TGCTAATCCC AGAACCTTTC CTTTTGCTTT TTTTAAGGAC ATTTGGGAAG TTCCTGGTGT 720

AGGACCCTTC TCCCTGGGAT AAGAAACCTG CCTGTAAACG CTCTGTAAAT ACTCCCTTCC 780

ACCCATCCCA GCCCCTGGGC AGCCGGGCAG AAGGGAATCC AGGCTATGGA CCTCCCAAGT 840

CCCCGCTCCC CGCTCCCCTC GGCGGCCCCG CCTTGTTCTG ATCTGTGTGT GAGTGTGTGT 900

GAACTTCTGA AAGACAATAT TAAAGAGACT TAGTTGAAAA AAAAAAAA 948

(2) INFORMATION FOR SEQ ID NO: 55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 990 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

GGGGAACTGC AGTGACAGCA GGAGTAAGAG TGGGAGGCAG GACAGAGCTG GGACACAGGT 60

ATGGAGAGGG GGTTCAGCGA GCCTAGAGAG GGCAGACTAT CAGGGTGCCG GCGGTGAGAA 120

TCCAGGGAGA GGAGCGGAAA CAGAAGAGGG GCAGAAGACC GGGGCACTTG TGGGTTGCAG 180

AGCCCCTCAG CCATGTTGGG AGCCAAGCCA CACTGGCTAC CAGGTCCCCT ACACAGTCCC 240

GGGCTGCCCT TGGTTCTGGT GCTTCTGGCC CTGGGGGCCG GGTGGGCCCA GGAGGGGTCA 300

GAGCCCGTCC TGCTGGAGGG GGAGTGCCTG GTGGTCTGTG AGCCTGGCCG AGCTGCTGCA 360

GGGGGGCCCG GGGGAGCAGC CCTGGGAGAG GCACCCCCTG GGCGAGTGGC ATTTGYTGCG 420

GTCCGAAGCC ACCACCATGA GCCAGCAGGG GAAACCGGCA ATGGCACCAG TGGGGCCATC 480

TACTTCGACC AGGTCCTGGT GAACGAGGGC GGTGGCTTTG ACCGGGCCTC TGGCTCCTTC 540

GTAGCCCCTG TCCGGGGTGT CTACAGCTTC CGGTTCCATG TGGTGAAGGT GTACAACCGC 600

CAAACTGTCC AGGTGAGCCT GATGCTGAAC ACGTGGCCTG TCATCTCAGC CTTTGCCAAT 660

GATCCTGACG TGACCCGGGA GGCAGCCACC AGCTCTGTGC TACTGCCCTT GGACCCTGGG 720

GACCGAGTGT CTCTGCGCCT GCGTCGGGGG NAATCTACTG GGTGGTTGGA AATACTCAAG 780

TTTCTCTGGC TTCCTCATCT TCCCTCTCTG AAGGACCCAA GTCTTTCAAG CACAAGAATC 840

CAGCCCCTGA CAACTTTCTT CTGCCCTCTC TTGCCCCANA AACAGCANAA GCAGGANANA 900

NACTCCCTCT GGCTCCTATC CCACCTCTTT GCATGGGAAC CTGTGCCAAA CACCCAAGTT 960

TAAGAAAAAA ATAAAACTGT GGCATCTCCA 990

(2) INFORMATION FOR SEQ ID NO: 56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1603 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

GGTCGACCCA CGCGTCCGGC CCGCCGGCTC CGGAGCGGCT CTGCCTTCCC GAGCGCGGGA 60

CCGCGCCCTG GGGGAGGAGG GCGAACGACG CGGCGATGGC TCCGCGGGCA CTCCCGGGGT 120

CCGCCGTCCT AGCCGCTGCT GTCTTCGTGG GAGGCGCCGT GAGTTCGCCG CTGGTGGCTC 180

CGGACAATGG GAGCAGCCGC ACATTGCACT CCAGAACAGA GACGACCCCG TCGCCCAGCA 240

ACGATACTGG GAATGGACAC CCAGAATATA TTGCATACGC GCTTGTCCCT GTGTTCTTTA 300

TCATGGGTCT CTTTGGCGTC CTCATTTNGC CAMCTNGCTT NAAGAAGAAA GGCTATCGTT 360

GTACAACAGA AGCAGAGCAA GATATCGAAG AAGAAAAAGG TTGAAAAGWT AGRATTGAAT 420

GACAGTGTGA ATGAAAACAG TGACACTGTT GGGCAAATCG TCCACTACAT CATGAAAAAT 480

GAAGCGAATG CTGATGTYTT AAAGGCGATG GTAGCAGATA ACAGCCTGTA TGATCCTGAA 540

AGCCCCGTGA CCCCCAGCAC ACCAGGGAGC CCGCCAGTGA GTCCTGGGCT TTGTCACCAG 600

GGGGGACGCC AGGGAAGCAC GTCTGTGGCC ATCATCTGCA TACGGTGGGC GGTGTWGTCG 660

AGAGGGATGT GTGTCATCGG TGTAGGCACA AGCGGTGGCA CTTTATAAAG CCCACTAACA 720

AGTCCAGAGA GAGCAGACCA CGGCGCCAAG GCGAGGTCAC GGTCCTTTCT GTTGGCAGAT 780

TTAGAGTNAC AAAAGTGGAG CACAAGTCAA ACCAGAAGGA ACGGAGAAGC CTGATGTCTG 840

TTAGTGGGGC TGAAACCGTC AATGGGGAGG TGCCGGCAAC ACCTGTGAAG AGAGAACGCA 900

GTGGCACAGA GTAGCAGGTG AGCCGTGGTT TTGGTGACAT TGGGGGCAGA GTGGTGCAGG 960

GTGAGGAGAA GGTACTTGGA GCCTCCCAGG TGCTGTGGCA GCATAGGAAT GGTATTTGAC 1020

AGGGAAGTGG GAGAGCTTTC CTTGACCCAG GAAGACTGAG GGGGACTGAA CATGATTACT 1080

TGTCTGCCTA GAGCTTCTTG TAAAGAAGTC ACAAACTTAG TGCCTCCAGG GGCTTGGCTG 1140

TGTGATAATG AGGATAGAGG ATTACTTGTG AGGCAATGTG GCATGGTGGG GATTGTGGCA 1200

AACTAGAATT CACATCACCC ACCATATAGG GCTTGCATTA CCACGAGGCA GAAAGCACCT 1260

AGTGTTGCTG CATCTTCTTA CGCAAAAAAG ACAAAATCCA GACTTCTAAA ATGTAAAATC 1320

ACTGATTTTC GATATTGGCA GCTTACTTTT TTTTTTTAAA CAACCATGCA GGCCAAATGA 1380

CTTGTAATCT TGTCACCATT TTTAGGTAAA CTGTGACTTG AAAAAGTCTG GAGCAAACAA 1440

ACCAATGCTT TTTCCTTTTA TTCTGTTGGR AACCAGTTTT CTTTGTGTCA CAGTTYTGAA 1500

ACCTCAATAC GAATATTTCT CTTCCCACCA AATATTTTGA GGCAATTGAA AAGCCACAGT 1560

GATTTATTTC TTGATTTGGC AATTTTAATT TTGCAAGACA ATT 1603

(2) INFORMATION FOR SEQ ID NO: 57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1052 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

TACAGCTCAG GATGCCTGTA ACATTGTCAT CTCTGGGCTT CTGGGTCCTG CTTAGCCTGC 60

TTTTTCCCTG GAGGACTGAC CAGGGATGCG GCCCAGCAAC ATGTTACTAA ATCATACTCT 120

CCTCCCTACC TTTCCCAGAC CTCTCACTCC TGCCTGGTGT TCCAACCCGT TCTGTGGCCA 180

GAGTATACAT TTTGGAACCT CTTCGAGGCC ATCCTGCAGT TCCAGATGAA CCATAGCGTG 240

CTTCAGCAGN AAGGCCCGAG ACATGTATGC AGAGGAGCGG AAGAGGCAGC AGCTGGAGAG 300

GGACCAGGCT ACAGTGACAG AGCAGCTGCT GCGAGAGGGG CTCCAAGCCA GTGGGGACGC 360

CCAGCTCCGA AGGACACGCT TGCACAAACT CTCGGCCAGA CGGGAAGAGC GAGTCCAAGG 420

CTTCCTGCAG GCCTTGGAAC TCAAGCGAGC TGACTGGCTG GCCCGTCTGG GCACTGCATC 480

AGCCTGAATG AGGCTGGCCA CCTGCCACTT TGCCCTGCCC TCTGCCTCCA GGGCTCCMCT 540

MYCCTTCCTT TTCTTGGTGA AAGGCACCTC CTTTCCTGAT AATGAATGGT GTTCCCTTTG 600

CTTGGCTGGG GAGCCCCCCA GGCCAGGTTT GCTGGCCATA GATACCTTTG GGCTGCCTGR 660

GACAGGCTCC TGAGGAGGAT TGAGGGTGAA AGTCTCCCAC GAGTACACTA AACCTAGGTC 720

TGGTCACCAA TAGGGTTTGG AGAGCAAAGG GCCACAACTC ATCAGCTGCC TGTCTCTTAG 780

ATGCACTTTC TTTTTCCACC AGCACATCCT TCAACACACA GAATTTCAGG GAAGAGTTCT 840

CCCCAAAACC CTAGCTCTTT ACCCTTCCAT TTTAGCCTTC CACCCAGCTT CCACAAAAGA 900

TTTGGCTCTA CCTTGGATCT GCTAGTAAAT AACTAATAGG CAGGCAGTTA TTTGGGTAAG 960

GAAAAAAGGG GTGGGAGAGA CAGAAAATTT GCCCACTGCT GCTCCTCCCC TTGGSTYTCC 1020

ACCTGGGATT TGCTATTGAA TCTCTACCCT NN 1052

(2) INFORMATION FOR SEQ ID NO: 58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 814 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

ACNCGNTGGC GGCCGCTCTA GAACTAGGGG ANCCCCCGGG CTGCAGGAAT TCGGCACGAG 60

CATAGACTTT TAAACTGGTA CGGTTCTTAG AGATGGTCCT TGGCCTTCTG TTGTTGTTGT 120

KGTTTTTTTC TTTTTCTTCT TCTCCTTCTC CTTCTTCTTC TCTTCTCCTT CTTTCTTCTT 180

TTTTTTTTCA GAGTCTTGCT CTGTCACCAA GACTGGAGTG AAGTGATGTG ATCTCGGCTT 240

ACTGCAACCT GGGAGGCAGA GGTTGCAGTG AGTCGAGATG GTGCCATTGC TCTCGTTTGG 300

GCAACAAGAG TGAAACTCTT GTCTCAAAAA AAAAAAAAAA ATGAGGTTTA AGACAGTTTT 360

GTCATTACTG GTGGGATCTG GTCACACAAG ATAGCATTAA ACGTGACATG GCACATAAAA 420

TTGGTTAAAA AATTTTGTTT TTTAATTACG TAATGTAAAA GCCCAACAAA CACTTTATGC 480

AAGATTGGAA TGTATCTTCA AATTCAGATT TAATAAACAT GTAAAGATCC TCTGTATATA 540

AAAGTTGTAT TTAATCCCTT GTGCCCCAAG AATGCTATAA AAGATCCCAA GAATGTTATC 600

TATGAAAAGA TAGCAATAGG GAATGGTGAA CAAATAATTT AATTTGCCAA TTCTAAAAAA 660

CATGGACTTA AACCCCATGA AAACTTGGTT CCATAGTTTT AACTGTTTTA TGGTTCCAAT 720

ACAAAACCAG AGTGGTTTAC ATTCCACAAT NACCAAATTT GCATCCAATN TTGGGGTAAT 780

TTTNGGTATT TGCCATGGGA TACTATTCAT TTTT 814

(2) INFORMATION FOR SEQ ID NO: 59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1215 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

AGAGGAAGTC TTTTGCCAAG CCTGTTCTCT GGACTAACGC CATCCAGGCT GGGAGGGGAA 60

GAGTGCTCTG CTACACTCGT CCCCCTCCTG CCTCATCTTC CTTCTCAGCC TTGGTTCCTG 120

ATGGGAACAG AATGGAGGGC CTGAGAACAT ACTTTCTAAA TGCCTTTGAC CCAGGAACCG 180

ATTATCTATA TTTGTTCCCA TTTTCCTTCA CCGTGACATT CCAGCATTGT CTGACTGTGA 240

GGTGGGCCTT TGAGAGCCTC CAGGTTCCTC AAAACAGGCC TGAGCGATGG GCATCACACC 300

CTCTGCCTAC CCACRTGCCT GCTTACCTGC CAGATAACCA AGTGNAGATG TCTGCGAGTG 360

GCTAGTTTTC ACATTCTTAC TAGTGTTTGG YTCACCTTTG GGCAAAGGCC CCCTCTAGGC 420

CTTGCCCCAC CTCCATCAAA CGCAGACACT GTAGTCAGAC CTCAGYAATA TAGGAGGCAA 480

TAATCTTTTA ACAGTGTTTT GCAAACAAAC AAAAAGAGAA AAATCCCAGC CAGGGGAACT 540

CGCCACCTGC CCACGCTAGT TCCATCCACG CTCAAGACCC GCCCTTAGAC CAGGCAGGCA 600

AAGGCCCCCA TCACACTCGG CCACTAGTGG GGTCCTGAGG CCAAGAAAGA AACCAGACCC 660

TGTATGACAA GTTGGGKTCT TTCCAGAACA CGACAGAAAC AGGGGGGGCC CCTTGTTAAT 720

GCCACTCCAT ACTCCAGAAG CATTATTCCT TATTTGGGAC AGCCAAGGGC AGATTCACAG 780

GTTATTGTAG GAATAAAGAC TAGTTTACAA AGGARAAAGA GSCCCTGGAC TTCCCMAGGA 840

AAGGTCAGGT TAGGGCTCCT GTACCCATTC TGTTCCACCA CTGTTTGATC TCTCTGGCCT 900

CCCACCAGGA ATGCCGTTTC CTTTTTATGG ATCTGTTGGG AACCAGAGAG AATCAACAGA 960

TCAATGACAT AGGATCCGAA GTGCAATGAT AGTCACTTCT AGTTTGGCAT TTCACAAACT 1020

CTGNACAGCA AGGTATTGGT AGGTTACTCA ATTTCAAAAG GGCCCCATGG CCAAATATGT 1080

TTAGGAACCG CTGTTTGNAT TTCTTTTTTT GGAGACGCAT TGTATATAAT ATATGTCAAA 1140

GGCTTTCGGA ATTCCTGCAG GAAAGAAATC AGCTTTGTTA AATCCNAAAA AAAAAAAAAA 1200

AAAAAAATAG ACTCG 1215

(2) INFORMATION FOR SEQ ID NO: 60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 478 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

ATTTCTTATG ACATGGGGGT TTGAATTGGT TGGCAAATGT TTAATTTTAA TATCCATAAT 60

CAGTGAGGTC CTGCTGGCTG TAATCATTAA TTGTGAAATC TAAGGAGCTT AGTTCATGGC 120

TCTAGAATTT CACAGAAAAR TGYGMTATGA TACGAGCATT AAGTTTATTT CTTCTGATCT 180

TTGATGCAGC TTTGTTCAGT TTATCTGTTT TTGTATTTAT TGGTCATCTA CTTCCCATGC 240

CAAAAGGGAC TGGTCTACAT AGCTGCGCTA AACACCTGAT CAAATCACTA AAAGAAAATG 300

TGTTACCTCT AATGAATTAT CCTGATTGTA AGTTAAAAAT CAATATTTCC CCGTAGTGAG 360

GTTTGCTTTT TAAAAAGAAK KCTTAAAAAA AAAAAAAAAA AAACGAGTTN AAGAAAAGGA 420

AGCAAGCTCA GGTAAGGTGC ACACATTGGG CTAAGGAAGC TAGAGCCTGT GGAGANGC 478

(2) INFORMATION FOR SEQ ID NO: 61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 618 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

TATGACCTTG ATAACCCCAA GTTNGAAATT AACCTTCANT AAAGGGAACA AAAGCTGGAG 60

TTCGCGCGCT TGCAGTTCGA CACTAGTGGA TCCCAAAGAA TTCGGCACGA GTCATAATGA 120

GCTACTAGGT AAGCCTTCTG GGACTTTCAG ATATTTTGGG GAAGATTGAT TTTTGTTCTT 180

ACATGCTGTG GACCCTTGGC CATCAAATGG TATGGGGAAG CTCATCCGTC TGTCTGTGAT 240

GGTCATGTCA GTCAGGCGTC TTTTTAGTAT TTACTGGGTG CTCAGTACTG TGCCAGATGC 300

TGTCGGGAGC CGTGGTGGTA TGGAGGAGGA GTGCTCCAGA GGACTCTGCT GTGTGGCAGG 360

CCAGCATAAA CAAGCCAAGG GGAAAAGGCA GGCATGGAAT AAAGGGGGAG AATACCAGTG 420

TGTGACTTAC TGCTGACTGT GTGGATTAGC CTATCAGCAG TAATCAAGCA GGGCGGAGGG 480

CATTATCTTT GAGCCAGAAG AGTGAGCACT GGSCCGAGGG TGGAGCATCA AGAGGGGGTG 540

TAGGACCNCA AGGCTTCTTN CNGGGGAGAC AACGTCAATA AGCNGTCAGT AGTCACCGAC 600

AGTTTTGGGA AGCAAGGG 618

(2) INFORMATION FOR SEQ ID NO: 62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 751 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

TCGACCCACG CGTCCGAGGA GCTGGACTTC TGAGACAGCC ATTCTCCTTG CATAGCACTG 60

TCTGCTGCTA CAGCTCATAG AAGTCAACAA TTTTCTTCAA CACTGGTAGG CAGCCTCTAA 120

ATGGCCCTGA TCACCCTCAC CTCCTGCCAT TCACACCNNT GTAAAATTCC ACCCCTGGAC 180

CTAGTGACTC ACTTCTAACA ANGAGAATAC AGCAAAAGTA ACATCGCTTC TGAGGTGAGG 240

CTACAAGGAG ACTACGATGC CTGCCTTGGT CACCCTTCTC CTGCTCTTTC CATTGCTCCC 300

TCTGATGGAA GCCAGTTGCC ATGTGATGAG GTGCCCTATG GAGAGGCCCA CGTGACAAGG 360

TATTGTAAAA AGCCTCTGAC CAATAGCCAT CTAGAAACGG AGGCCCAGTC CAGCAGCCTC 420

TGAGATGAAT CCTGCCAACC TGAGCTTGGA GACAGATTCT CTCCCTATCC TGCCTTGGGA 480

TGATCACAGC CACCACCAAC ACCTTCACTG CCTGGTGAGA GGCCAAGCCA GTGAACCCAA 540

GGTAAACTGG ACAGAATCCT GACCCACAGA AACTGAGATA ATGTTTGTTA TTTTAAGCTG 600

CTCAGTTTGT TACAGAGCAA TAGATAACTA ACTCAAACAC CATAAAATTC TAATATTTTA 660

TTCTATCACA CAAACCAGGT AATACCAAGT AAATGCCATT ACTATACACA TATTTTTGTA 720

ACACAATTAC ATGTGATTTT TTAAGAAGGC T 751

(2) INFORMATION FOR SEQ ID NO: 63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 780 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

CNGNCAGTCA CNGTCCCCGA TTCCCGGGTC GACCCACGCG TCCGGGTTGG CAACTCCTGA 60

GGCCTGCATG GGTGACTTCA CATTTTCCTA CCTCTCCTTC TAATCTCTTC TAGAGCACCT 120

GCTATCCCCA ACTTCTAGAC CTGCTCCAAA CTAGTGACTA GGATAGAATT TGATCCCCTA 180

ACTCACTGTC TGCGGTGCTC ATTGCTGCTA ACAGCATTGC CTGTGCTCTC CTCTCAGGGG 240

CAGCATGCTA ACGGGGCGAC GTCCTAATCC AACTGGGAGA AGCCTCAGTG GTGGAATTCC 300

AGGCACTGTG ACTGTCAAGC TGGCAAGGGC CAGGATTGGG GGAATGGAGC TGGGGCTTAG 360

CTGGGAGGTG GTCTGAAGCA GACAGGGAAT GGGAGAGGAG GATGGGAAGT AGACAGTGGC 420

TGGTATGGCT CTGAGGCTCC CTGGGGCCTG CTCAAGCTCC TCCTGCTCCT TGCTGTTTTC 480

TGATGATTTG GGGGCTTGGG AGTCCCTTTG TCCTCATCTG AGACTGAAAT GTGGGGATCC 540

AGGATGGCCT TCCTTCCTCT TACCCTTCCT CCCTCAGCCT GCAACCTCTA TCCTGGAACC 600

TGTCCTCCCT TTCTCCCCAA CTATGCATCT GTTGTCTGCT CCTCTGCAAA GGCCAGCCAG 660

CTTGGGAGCA GCAGAGAAAT AAACAGCATT TCTGATGCCA AAAAAAAAAA AAAAAAAACC 720

GCGGCCGAAA GCTTACTNCC CTTTAAGTAA GGGGTTAATT TTTAGCTTGG GCACTNGGCC 780

(2) INFORMATION FOR SEQ ID NO: 64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 588 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

TTCCGAATTA ATCGACTCAC TATAGGAAWT GCCGTCGCCA TGACCCGCGG TAACCAGCGT 60

GAGCTCGCCC GCCAGAAGAA TATGAAAAAG CAGAGCGACT CGGTTAAGGG AAAGCGCCGA 120

GATGACGGGC TTTCTGCTGC CGCCCGCAAG CAGAGGGACT CGGAGATCAT GCAGCAGAAG 180

CAGAAAAAGG CAAACGAGAA GAAGGAGGAA CCCAAGTAGC TTTGTGGCTT CGTGTCCAAC 240

CCTCTTGCCC TTCGCCTGTG TGCCTGGAGC CAGTCCCACC ACGCTCGCGT TTCCTCCTGT 300

AGTGCTCACA GGTCCCAGCA CCGATGGCAT TCCCTTTGCC CTGAGTCTGC AGCGGGTCCC 360

TTTTGTGCTT CCTTCCCCTC AGGTAGCCTC TCTCCCCCTG GGCCACTCCC GGGGGTGAGG 420

GGGTTACCCC TTCCCAGTGT TTTTTATTCC TGTGGGGCTC ACCCCAAAGT ATTAAAAGTA 480

GCTTTGTAAT TCCAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 540

AAAAAAAAAA AAAAAAAAAA AAAANNCGGG GGGGGGCCCC CCCCCCCC 588

(2) INFORMATION FOR SEQ ID NO: 65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 774 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

TTTAAAGATG AAGAAATGAC AAGGGAGGGA GATGAGATGG AAAGGTGTTT GGAAGAGATA 60

AGGGGTCTRA GAAAGAAATT TAGGGCTCTG CATTCTAACC ATAGGCATTC TCGGGACCGT 120

CCTTATCCCA TTTAATTAAT TTCTCTGACA ATTCAATTAT TTTCTGTTAT TAATGTTGCC 180

ACTGCTTTCT GTTTGTCTGC ACTTTCTTGA TAAATATTTG CTATCGTTTT ACTCCAGTCA 240

TTCGATGTTG CTGAGATTTA CATATGACTC TTGTCAACAT CTCATCTTTT GACCCAATCT 300

TATTCATTTA ATAAGAGGTC TCATTCATTT GCATGGAAAA ATGCTCATTG TATATTGCAA 360

AGTGAAAATA ACGAGTTGCA AAACAGTGTA TACATATATG TGTGTATATA TGTACACTTT 420

ATTTGTACAT TTCTATGTGA CATAATGCAA AGGAAAGTGT CTGATTTTAT TATACACCAA 480

AGGTTAACAG TGAATCTCTG TGTGATCTCT TTTTTTTTCT TTTTGCCTAT CTGCATCTTC 540

TCACTTGCCA AAAAATGAAT ATATGTTTAT GTGTGTATAT TACTTGTGTC ACAAAAAACC 600

CTAAAGTAGA CAGTAAAAGA ACTTGTCAAT CGCCTTTGGA AGGCAATGAA ACACTTAATA 660

AACTCTCAAT AACAGAAGCG TAAAAATGAA ATGTAAACCT CCAATTACCT CTGGATCTCT 720

TAGCCAGAGT AATAAACTGG TAATTATTAC AGATAAAAAA AAAAAAAAAA AANA 774

(2) INFORMATION FOR SEQ ID NO: 66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1866 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

ACCCACGCGT CCGGTCCTCT TCTTCAGCAC ATGCCAAAGC TGTTCCTCAC GGCCTGTGAG 60

ACAAGAGCAT CTTGGATGTA GGACAATGGA AGAGTTAGAT GCCTTATTGG AGGAACTGGA 120

ACGCTCCACC CTTCAGGACA GTGATGAATA TTCCAACCCA GCTCCTCTTC CCCTGGATCA 180

GCATTCCAGA AAGGAGACTA ACCTTGATGA GACTTCGGAG ATCCTTTCTA TTCAGGATAA 240

CACAAGTCCC TTGCCGGCGC ANTCGTGTAT ACTACCAATA TCCAGGAGCT CAATGTCTAC 300

AGTGAAGCCC AAGAGCCAAA GGAATCACCA CCACCTTCTA AAACGTCAGC AGCTGCTCAG 360

TTGGATGAGC TCATGGCTCA CCTGACTGAG ATGCAGGCCA AGGTTGCAGT GAGAGCAGAT 420

GCTGGCAAGA AGCACTTACC AGACAAGCAG GATCACAAGG CCTCCCTGGA CTCAATGCTT 480

GGGGGTCTSG AGCAGGAATT GCAGGACCTT GGCATTGCCA CAGTGCCCAA GGGCCATTGT 540

GCATCCTGCC AGAAACCGAT TGCTGGGAAG GTGATCCATG CTCTAGGGCA ATCATGGCAT 600

CCTGAGCATT TTGTCTGTAC TCATTGCAAA GAAGAGATTG GCTCCAGTCC CTTCTTTGAG 660

CGGAGTGGCT TGGNCTACTG CCCCAACGAC TACCACCAAC TTTTTTCTCC ACGCTGTGCT 720

TACTGCGCTG CTCCCATCCT GGATAAAGTG CTGACAGCAA TGAACCAGAC CTGGCACCCA 780

GAGCACTTCT TCTGCTCTCA CTGCGGAGAG GTGTTTGGTG CAGAAGGCTT TCATGAGAAG 840

GACAAGAAGC CATATTGCCG AAAGGATTTC TTAGCCATGT TCTCACCCAA GTGTGGTGGC 900

TGCAATCGCC CAGTGTTGGA AAACTACCTT TCAGCCATGG ACACTGTCTG GCACCCAGAG 960

TGCTTTGTTT GTGGGGACTG CTTCACCAGT TTTTCTACTG GCTCCTTCTT TGAACTGGAT 1020

GGACGTCCAT TCTGTGAGCT CCATTACCAT CACCGCCGGG GAACGCTCTG CCATGGGTGT 1080

GGGCAGCCCA TCACTGGCCG TTGTATCAGT GCCATGGGGT ACAAGTTCCA TCCTGAGCAC 1140

TTTGTGTGTG CTTTCTGCCT GACACAGTTG TCGAAGGGCA TTTTCAGGGA GCAGAATGAC 1200

AAGACCTATT GTCAACCTTG CTTCAATAAG CTCTTCCCAC TGTAATGCCA ACTGATCCAT 1260

AGCCTCTTCA GATTCCTTAT AAAATTTAAA CCAAGAGAGG AGAGGAAAGG GTAAATTTTC 1320

TGTTACTGAC CTTCTGCTTA ATAGTCTTAT AGAAAAAGGA AAGGTGATGA GCAAATAAAG 1380

GAACTTCTAG ACTTTACATG ACTAGGCTGA TAATCTTATT TTTTAGGCTT CTATACAGTT 1440

AATTCTATAA ATTCTCTTTC TCCCTCTCTT CTCCAATCAA GCACTTGGAG TTAGATCTAG 1500

GTCCTTCTAT CTCGTCCCTC TACAGATGTA TTTTCCACTT GCATAATTCA TGCCAACACT 1560

GGTTTTCTTA GGTTTCTCCA TTTTCACCTC TAGTGATGGC CCTACTCATA TCTTCTCTAA 1620

TTTGGTCCTG ATACTTGTTT CTTTTCACGT TTTCCCATTT CCCTGTGGCT CACTGTCTTA 1680

CAATCACTGC TGTGGAATCA TGATACCACT TTTAGCTCTT TGCATCTTCC TTCAGTGTAT 1740

TTTTGTTTTT CAAGAGGAAG TAGATTTTAA CTGGACAACT TTGAGTACTG ACATCATTGA 1800

TAAATAAACT GGCTTGTGGT TTCAATAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1860

AAAAAA 1866

(2) INFORMATION FOR SEQ ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1152 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

CTCAAGGATG TAAAGGCTCT GCAGATTTCG GGAGGCCTGT CTCCCAGCAC CTGATGGGAC 60

ACTTTTTGCC CCACTGTAAA TTCTGGGTGT ATCCTCCACT GTATGCTGTC ACCCCAAGGG 120

CAAGCACTGC ATCTGCTTAG TGAAGGATTT ATTGTTCGGA AGATACATTT TCCCCTTKAG 180

CAGAGAGTGG CGTATCCTGG CAGTCTTCGG TGAGCCAGTT GTACCAGGAT TATGAAATGC 240

AGATGTTTAC TGTGTCATTG TTGCTGTCAT TGCTACTGAG GAGTACTGAC CAGAATCATC 300

TGCAACTYTT AGTTGGCAGA GAGGACCACT ATGGCGGGTA GCTCTTTTCT TTCCTGCCAT 360

TGTGGGGATG ATTCCAGGCC AAAGATGATG GARAAGTATG GAAATCATCT GAAAGGTTGA 420

AGCTTGGCAC GTGAAGCCAT TCATGACTTT GTAAGGCAGT TTTGCTGAAG GCCAGTTCTG 480

CCCTGGGAGG GACGGAGGTG AATCCTCCTG AGTACCTGTG GTTTTCTTAC TTCCTGCTGA 540

ATTTACCTAA GTGCCTGTTG TTTGCTTGCT GTGGAGGCTT TCTGGTATTT CATTTCAGGT 600

GCAGATGCCT TCACTTTCCC ACCRAAAAAA CCCCMACCAA ACCTAAGACC TTACTGCAAC 660

TAAGTYTNCC AAGTACTTTT TAACCCAATG GGATGAACAG CCTGTGGTCT GCTCAGATCA 720

CCCTGAGTGC GTGTGAGAAG GCMTNGGCTT TGCCAGGAAA TCCAGGAAGG CAGGGCCGGG 780

CTGTGTTGGA AGCTGGCTTA GCTGGTGGGG CAGCCTTATT TCAATTAAAA GGGCATTGAC 840

TGGGAGCAGC AGTCCTGGAG TTTGTTGCAT TTCCTATTGC CCTCAAAATG AGAAACCAGG 900

AAAATAGCAG ATTGGAGCCT TCGAGAAGGC AGTAAATGGC TGTTTTTATT GACAAAAGGA 960

AAACATTTTA CTGCCATCTC ACTGATGGCA TCTCACTGAC TTAAAATGAA GGCANGTTGT 1020

AGTAAAAAAA AAAGTCTACA TTTTTCCACC GCCACGTTCT TATATCCTGT TTGTCAGCCA 1080

CTGCTCANAA GGGCATGTTG TCTTGCGGAN TANAGGCGCT CTCCTTCCCT CGTTTTCCCT 1140

ATAGGTTGGG TG 1152

(2) INFORMATION FOR SEQ ID NO: 68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2483 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

AGCAGGCGGT GCGCTGGGGG CGGGAGCAGC GCGKAGCCCG GCTCGGCCAC ACCGATCGCC 60

CGCCGCCATG GGCTCCTCGC AAAGCGTCGA GATCCCGGGC GGGGGCACCG AGGGCTACCA 120

CGTTCTGCGG GTACAAGAAA ATTCCCCAGG ACACAGAGCT GGTTTGGAGC CTTTCTTTGA 180

TTTTATTGTT TCTATTAATG GTTCAAGATT AAATAAAGAC AATGACACTC TTAAGGATCT 240

GCTGAAASCA AACGTTGAAA AGCCTGTAAA GATGCTTATC TATAGCAGCA AAACATTGGA 300

ACTGCGAGAG ACCTCAGTCA CACCAAGTAA CCTGTGGGGC GGCCAGGGCT TATTGGGAGT 360

GAGCATTCGT TTCTGCAGCT TTGATGGGGC AAATGAAAAT GTTTGGCATG TGCTGGAGGT 420

GGAATCAAAT TCTCCTGCAG CACTGGCAGG TCTTAGACCA CACAGTGATT ATATAATTGG 480

AGCAGATACA GTCATGAATG AGTCTGAAGA TCTATTCAGC CTTATCGAAA CACATGAAGC 540

AAAACCATTG AAACTGTATG TGTACAACAC AGACACTGAT AACTGTCGAG AAGTGATTAT 600

TACACCAAAT TCTGCATGGG GTGGAGAAGG CAGCCTAGGA TGTGGCATTG GATATGGTTA 660

TTTGCATCGA ATACCTACAC GCCCATTTGA GGAAGGAAAG AAAATTTCTC TTCCAGGACA 720

AATGGCTGGT ACACCTATTA CACGTCTTAA AGATGGGTTT ACAGAGGTCC AGCTGTCCTC 780

AGTTAATCCC CCGTCTTTGT CACCACCAGG AACTACAGGA ATTGAACAGA GTCTGACTGG 840

ACTTTCTATT AGCTCAACTC CACCAGCTGT CAGTAGTGTT CTCAGTACAG GTGTACCAAC 900

AGTACCGTTA TTGCCACCAC AAGTAAACCA GTCCCTCACT TCTGTGCCAC CAATGAATCC 960

AGCTACTACA TTACCAGGTC TGATGCCTTT ACCAGCAGGA CTGCCCAACC TCCCCAACCT 1020

CAACCTCAAC CTCCCAGCAC CACACATCAT GCCAGGGGTT GGCTTACCAG AACTTGTAAA 1080

CCCAGGTCTG CCACCTCTTC CTTCCATGCC TCCCCGAAAC TTACCTGGCA TTGCACCTCT 1140

CCCCCTGCCA TCCGAGTTCC TCCCGTCATT CCCCTTGGTT CCAGAGAGCT CTTCTGCAGC 1200

AAGCTCAGGA GAGCTGCTGT CTTCCCTCCC GCCCACCAGC AACGCACCCT CTGACCCTGC 1260

CACAACTACT GCAAAGGCAG ACGCTGCCTC CTCACTCACT GTGGATGTGA CGCCCCCCAC 1320

TGCCAAGGCC CCCACCACCG TTGAGGACAG AGTCGGCGAC TCCACCCCAG TCAGCGAGAA 1380

GCCTGTTTCT GCGGCTGTGG ATGCCAATGC TTCTGAGTCA CCTTAACTTT GAACCATTCT 1440

TTGGAATTGG CGTGGTATAT TTAACCACGG GAGCGTGTCT GGAAACGCAA ACTATCATTA 1500

ATTTCATACT AGTTTGTACC GTATCTGTAG GCATCCTGTA AATAATTCCA AGGGGAAAAC 1560

TAAACGAGGA CGTGGGTTGT ATCCTGCCAG GTTGAGTGGG GCTCACACGC TAGGGTGAGA 1620

TGTCAGAAAG CGCTTGTATT TTAAACAACC AAAAAGAATT GTAAGGGTGG CTTGCTGCCA 1680

GGCTTGCACT GCCGTTCCTG GGGGTGTGCA TCTTCGGGAA AGGTGGTGGC GGGGCGTCCA 1740

CTAGGTTTCC TGTCCCCTGC TGCTCCTTCC GTAAGAAAAT GAAATATTCT ATGCCTAATA 1800

CTCACACGCA ACATTTCTTG TACTTTGTAA GTCGTTTGCG AGAATGCAGA CCACCTCACT 1860

AAACTGTAAA CGGTAAAGAG ATTTTTACTT TTGGTCTCCG TGAGTCGCAT CTCTACTAAG 1920

GTTTACACAG GAATTCCACC TGAAGACTTG TGTTAAAGTT CTACAGCGCG CACTGTTAAC 1980

TGAACGTCTT TTTCTTCAGC CTATACGCGG ATCCTTGTTT TGAGCTCTCA GAATCACTCA 2040

GACAACATTT TGTAACTGCT GCTGTTGCTT TCTACATACA CCTTATAAAG TGACATTTCA 2100

AAAGAAATAA GGTGCCACAG TTTTAAACCA GAAGGTGGCA CTCTGTGGCT CCTTGTAGTA 2160

TTATAGCTAT ACTGGGAAAG CATAGATACA GCAATAAAGT ACAGTAATTT TACTTTTTTT 2220

CTTGTGTTAC ATCTAAATTA CAACCCTTAA TTGCCACGTG TGCACTTACT ACTCTCCAGT 2280

ATGTCTTATT ACTCTCCAGT ATGTCACGCA TCTTTAACTT TTCACGTCCT ATGTTTGCTT 2340

TCTCCCATTT TTAAGAGATG GTAAGTTAAC TGGAATTGAT TTACTGAATG AAATTAAATG 2400

CAGATATCCC TGTTTTTGAA ATAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 2460

AAAAAAAAAA AAAAAAAAAA AAA 2483

(2) INFORMATION FOR SEQ ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 536 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

GAGAAATGGA GCTTTGTTAG ATAAAAATTT TTTCAACGCA AACAGTCATT TTCCAGTGAA 60

AGGAGAGCGT ATCCGCCGTA GGATGGACTT AGATCGTGTA AAAGCTGAGG CCACCGAGGA 120

TATAACCTCC GGGGTCCTTT GCCTCCTTTT CCTTAGACTC CCTCCAAACT CGTGTATCTT 180

TCCTTCAGCA GTACTGGGCT CCACGCGAAC CTAGTCCTTT GTCTTTACCC TATTACCTTT 240

CATAACATCC TAGTTGAAAA GTARTTATTC AACCGCGTTT GAAAATGAGA ACAGGTTCAC 300

AGARGCTAGG TTACTTGCGA AGGTCGTTCA ATTAGTAACC AGTAACGCCA GGACTGCCAG 360

TTTCTTGCTT CCGAATTCTC ATGGTAGCTT TCACCARGCT CCCCGTCMAA TGCTAACGTC 420

AACTACTGAA CTAGATTAGC AAAAAGGTCT TTTAACAGAA TTCCTGGTTT TCAGAGAGAG 480

TTTCTTTCAT GAAGCGCCCC ATTTCTACAG AGGAAAATAA ACTCCAAGCA GCCAGT 536

(2) INFORMATION FOR SEQ ID NO: 70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 865 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

CCACGCGTCC GGCCTTTCTT GGCCAGAGGC GCCGGTTGGA CTCACGGGCG GGGCATGATG 60

GGTAACAGGA CCGGTGGGGT CCCCAGGAAG TCCTAGAGGG GGTCGGGGTT TGGGTGGACA 120

AGCTTTCCTC GTCCTCTCCC GACAGAGCTG ACGTGTCCTG GGTTCCACCG GGAGCGGGCA 180

TTTCCACCGG ACGGGAGGGT TCGGGGTGTC CGGGGCTGGG GAATACGTAG GGGTTGCCGC 240

GCGGTGTGGG GAGTTGGGGC GTGTGGCTGC AGTCCCGGGA GTTCTTGGAG GGGGTCGGCC 300

CACCGAGCTT CCGGACCGGC TGATCTGCCC GTAGCTTGCC GGANGGARGG CGGAGCTGAC 360

TCTCCGTCCC TTCTCCCATC CCCTCCAGTG GTGGGTACGG GCACCTCGCT GGCGCTCTCC 420

TCCCTCCTGT CCCTGCTGCT CTTTGCTGGG ATGCAGATGT ACAGCCGTCA GCTGGCCTCC 480

ACCGAGTGGC TCACCATCCA GGGCGGCCTG CTTGGTTCGG GTCTCTTCGT GTTCTCGCTC 540

ACTGCCTTCA ATAATCTGGA GAATCTTGTC TTTGGCAAAG GATTCCAAGC AAAGATCTTC 600

CCTGAGATTC TCCTGTGCCT CCTGTTGGCT CTCTTTGCAT CTGGCCTCAT CCACCGAGTC 660

TGTGTCACCA CCTGCTTCAT CTTCTCCATG GTTGGTCTGT ACTACATCAA CAAGATCTCC 720

TCCACCCTGT ACCAGGCAGC AGCTCCAGTC CTCACACCAG CCAAGGTCAC AGGCAAGAGC 780

AAGAAGAGAA ACTGACCCTG AATGTTCAAT AAAGTTGATT CTTTGTAAAA AAAAAAAAAA 840

AAAAAAAAAA AAAAAAAAAA AAAAA 865

(2) INFORMATION FOR SEQ ID NO: 71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 932 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

TCATCATATA CAAAGTTTTT CGTCACACTG CAGGGTTGAA ACCAGAAGTT AGTTGCTTTG 60

AGAACATAAG GTCTTGTGCA AGAGGAGCCC TCGCTCTTCT GTTCCTTCTC GGCACCACCT 120

GGATCTTTGG GGTTCTCCAT GTTGTGCACG CATCAGTGGT TACAGCTTAC CTCTTCACAG 180

TCAGCAATGC TTTCCAGGGG ATGTTCATTT TTTTATTCCT GTGTGTTTTA TCTAGAAAGA 240

TTCAAGAAGA ATATTACAGA TTGTTCAAAA ATGTCCCCTG TTGTTTTGGA TGTTTAAGGT 300

AAACATAGAG AATGGTGGAT AATTACAACT GCACAAAAAT AAAAATTCCA AGCTGTGGAT 360

GACCAATGTA TAAAAATGAC TCATCAAATT ATCCAATTAT TAACTACTAG ACAAAAAGTA 420

TTTTAAATCA GTTTTTCTGT TTATGCTATA GGAACTGTAG ATAATAAGGT AAAATTATGT 480

ATCATATAGA TATACTATGT TTTTCTATGT GAAATAGTTC TGTCAAAAAT AGTATTGCAG 540

ATATTTGGAA AGTAATTGGT TTCTCAGGAG TGATATCACT GCACCCAAGG AAAGATTTTC 600

TTTCTAACAC GAGAAGTATA TGAATGTCCT GAAGGAAACC ACTGGCTTGA TATTTCTGTG 660

ACTCGTGTTG CCTTTGAAAC TAGTCCCCTA CCACCTCGGT AATGAGCTCC ATTACAGAAA 720

GTGGAACATA AGAGAATGAA GGGGCAGAAT ATCAAACAGT GAAAAGGGAA TGATAAGATG 780

TATTTTGAAT GAACTGTTTT TTCTGTAGAC TAGCTGAGAA ATTGTTGACA TAAAATAAAG 840

AATTGAAGAA ACACATTTTA CCATTTAAAA AAAAAAAAAA ACTNGAGGGG GGCCCGGTAC 900

CCAAATCGCC GCATAGTGAT CGTAAACAAT CT 932

(2) INFORMATION FOR SEQ ID NO: 72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 996 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

CGCCTGGCAC CATGAGGACG CCTGGGCCTC TGCCTGTGCT GCTGCTGCTC CTGGCGGGAG 60

CCCCCGCCGC GCGGCCCACT CCCCCGACCT GCTACTCCCG CATGCGGGCC CTGAGCCAGG 120

AGATCACCCG CGACTTCAAC CTCCTGCAGG TCTCGGAGCC CTCGGAGCCA TGTGTGAGAT 180

ACCTGCCCAG GCTGTACCTG GACATACACA ATTACTGTGT GCTGGACAAG CTGCGGGACT 240

TTGTGGCCTC GCCCCCGTGT TGGAAAGTGG CCCAGGTAGA TTCCTTGAAG GACAAAGCAC 300

GGAAGCTGTA CACCATCATG AACTCGTTCT GCAGGAGAGA TTTGGTATTC CTGTTGGATG 360

ACTGCAATGC CTTGGAATAC CCAATCCCAG TGACTACGGT CCTGCCAGAT CGTCAGCGCT 420

AAGGGAACTG AGACCAGAGA AAGAACCCAA GAGAACTAAA GTTATGTCAG CTACCCAGAC 480

TTAATGGGCC AGAGCCATGA CCCTCACAGG TCTTGTGTTA GTTGTATCTG AAACTGTTAT 540

GTATCTCTCT ACCTTCTGGA AAACAGGGCT GGTATTCCTA CCCNGGAACC TCCTTTGAGC 600

ATAGAGTTAG CAACCATGCT TCTCATTCCC TTGACTCATG TCTTGCCAGG ATGGTTAGAT 660

ACACAGCATG TTGATTTGGT CACCTAAAAA GAAGAAAAGG ACTAACAAGC TTCACTTTTA 720

TGAACAACTA TTTTGAGAAC ATGCACAATA GTATGTTTTT ATTACTGGTT TAATGGAGTA 780

ATGGTACTTT TATTCTTTCT TGATAGAAAC CTGCTTACAT TTAACCAAGC TTCTATTATG 840

CCTTTTTCTA ACACAGACTT TCTTCACTGT CTTTCATTTA AAAAGAAATT AATGCTCTTA 900

AGATATATAT TTTAYGTAGT GCTGACAGGA CCCACTCTTT CATTGAAAGG TGATGAAAAT 960

CAAATAAAGA ATCTCTTCAC ATGARAAAAA AAAAAA 996

(2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 785 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

GGCACGAGGG GCTTTGCGTA CACAATAGCT GCTAGGAGTA CCCAAAGCCT GARTACARCC 60

TGCTGGTGTC ATGGCCACGT GTGAGCAGGC CAGCGTCAMA CGGCTCGCTG TGACCCGTCC 120

CGRAGACTGA AATGGGCCTG GGTCTTCTCC TKGTCCTGTG ATWAAAGTCC TCTCTTGAAA 180

GTGGAGAGCA AAGGCACACA GAGGTGCGCG CTCACAAGAA TTCCTCCCGG TGACTGGGTA 240

ATCAATGTTA CTGCTGTTTC CTTTGCAGGA AAGACCACAG CAAGATTCTT TCATTCGTCT 300

CCTCCTAGCC TGGGGGACCA GGCTCGAACT GACCCTGGAC ATCAAAGGAG GGATTATGTG 360

GCTGCTAAAG CCATCGGCCC ACAGCCCTGT TCACRTCTTG GTGCTTCTCT TTCCCAGAGG 420

CTGGTCCCAG CCAGGCACAC ACAAAAGGCA GATTCTCGTA AACSCAGCCT CCCTCCCTGG 480

AGGCTGCCTC CTGCCCTGGA TCTGGAGTGG AGCTGCTCTG AGATTTTGAG TTCTTCTGCA 540

GAGATGATTA AATATATCCA AGAGACATTG GAAAACCTGC TGAACATTTT ACATTGGTCT 600

GCTCAGCACA TGGCTGGATG CGGATATTTC TATAATTCCA GAAAGTCACA CAGCTCCTCT 660

GTATGAGACC AGTGGGCGCC ATTTAAAAGA ACAGGATGAG AATCTAAGAT ATATTATTAA 720

TAAATGTAAT GGATTTTTTT TTTGTAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 780

AAAAA 785

(2) INFORMATION FOR SEQ ID NO: 74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1069 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

TCCTCACCAT TCCCCTAGGN CAGGTCCCTG CAGGTCCCAC ACTTCTCCCA GGTCCCTAAA 60

CTTGGGTCGG TCCTTTCCCT GGAGTAGCTG GNTCCTCCAG TCGAGGTCCC TGTTCAGTCG 120

GTTCTTAGGC TCCTGCACAT GAAGGTGTGT GCCTGTGGTG TGTGGGCTGC TCTAGGAGCA 180

GATACAGGCT GGTATAGAGG ATGCAGAAAG GTAGGGCAGT ATGTTTAAGT CCAGACTTGG 240

CACATGGCTA GGGATACTGC TCACTAGCTG TGGAGGTCCT CAGGAGTGGA GAGAATGAGT 300

AGGAGGGCAG AAGCTTCCAT TTTTGTCCTT CCTAAGACCC TGTTATTTGT GTTATTTCCT 360

GCCTTTCCGA GTCCTGCAGT GGGCTGCCCT GTACCCTGAA CCTCATGAGC CTCTAAGGGA 420

AAGGAGGAAC AATTAGGACG TGGCAATGAG ACCTGGCAGG GCAGARTACA AGCCCAGCAC 480

CAGTGTCCCA GCCTTACTGG GTCCTTACCC TGGGCCAAAC AGGGAGGGCT GATACCTCCT 540

TGCTCTTCCT AGATGCCCAC CTCCTACAAT CTCAGCCCAC AAGTCCTCTC CACCCTAGGG 600

GGCTTGCTGC ATGGCAATAA CTCATAATCT GATTTGGAGG TTTGCCCTTT ACAGGGGCAG 660

ATTTTCTGCT CAGTTCAACA ATGAAATGAA GAGGAACTCC CTCTTTCTAC AGCTCACTTC 720

TATCAGAGGC CCAGGTGCCT CAGAGCCACA TTGAGTTGCT TTTTCTGGGA TGAGGAAGTA 780

GGGTTAAACT CCCCAGTTTC CTGAGGGAGG CTCCTGACAG GTGCCCTTTG TCAGACCCTA 840

CCACAGCCTG GATAGGCAGC CACATTGGTC CTCGCCCTTG CTCGGNACTC CGTGGTGGTC 900

CTGCCCTTCT CCCTGCATGC CTGTGGGTCT GCTCTGGTGT GTGAAGGTCG GTGGGTTAAC 960

TGTGTGCCTA CTGAACCTGG CAAATAAACA TCACCCTGCA AAGCCAAAAA AAAAAAAAAA 1020

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAA 1069

(2) INFORMATION FOR SEQ ID NO: 75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 831 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

GGACATTAGA TCACTGTGGA CCTAAAACAA ACAAACAACT ATAAGGAAAA TGGCATTAGA 60

AATGGTCTGG GGATCAGTTT ATCACTGCAG TTGTTACATC ACCCCATGGT CTAAAATACA 120

GAGCTTTAGT CTGTCTCTGT TTCAGTTCAT TTTACAGGAG GTGAACATCA CACTTCCAGA 180

AAACTCTGTC TGGTATGAAA GGTATAAATT TGATATTCCT GTCTTTCACT TGAATGGCCA 240

GTTTCTGATG ATGCATCGAG TAAACACCTC AAAACTTGAA AAACAGCTCC TGAAACTTGA 300

GCAGCAAAGT ACTGGARGCT GACTGATGCC CTCATGATTT TCCACCCTCT CTTCCCATAA 360

AGCATCTTCC TAAGGAAATG AMCATGGCCT GATACTCATT TTGTCACTTG TACAGAGCCC 420

TAAGGATGTT CTGAATTCAG TGGTGCCAAA TAAATGTTGA CATTCCCCTT TTGGTTGATG 480

GAAGTATCAG TGTGGGAACT GTTTGCTTAA TGGCATTTTA TAAAATAAKA AKAKCATATT 540

AGCAGGGAGG GAGATGATGG AGGGAGGGAG AAGTCCATTT GTCTTATTTA TCCTTTTTGT 600

ATTAATAGAG AAGCACTTCA CAGTCACTGG CAATGCCATT TATAGGAAGA AGGTTCTGCA 660

TTCCTGCTGC TCCCGGAGGG CTTAACTTTT TAATGAAAGA ATAAATGCTC TTCCACTCAG 720

TAGATAAAGT GAAATGTGAA TTGTTAATAA CTGTGCACGG TCAATAAAGC GATGTTTTAA 780

GGAATACAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAACTCG A 831

(2) INFORMATION FOR SEQ ID NO: 76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 590 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

TATATATAGA CNGTTAATAG TCGTGANTGN TGTGNACGAA CATTAACGGA AGTAGCATGT 60

AGCCAGTCGA ATAACNTATA AGGACAAAGT GGAGTCCACG CGTGCGGCCG TCTAGACTAG 120

TGGATCCCCC GGCTGCAGGA TTCGGCACGA GCTGCCAGGT GAGGAGCAGA GAGACTGTTC 180

CCTTGGGTGG AGAGGTGTGG GCATGAGAGC CACCCATTGC CAAGCAGCAA GAATGTTCGT 240

GCTTTTTTCC CTTCCAAAAT ATGCAGGGCT CAGGCTCCCA ATTCCGGGCC TGTCTGCTTT 300

GCTTGTGTTT CTCCTGTCCC TGTTCTCCCG GAGGGCCCAG GTGGAACTCA CGACAGGGAG 360

GGAGACGCTT CCCAAAAACC TGCAGGGCTA TTTCCCAGAA TTTGGTTTTC AAGTACAAAA 420

CTTTTTGTCC TGTAAGATAT ATGCAGCCTC ACAGAAGCAG CCTCTGCCTC CACTTTACCA 480

GCTACGTTTT TATCTTAAGC ACATGGGGCT CCCTTAGAAC TTACTCCACT GATTTAAAAA 540

AAAAAAAAAA AAACTCGAGG GGGGGCCCGG TACCCATTCG CCCTAAAAGT 590

(2) INFORMATION FOR SEQ ID NO: 77:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1274 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77:

GAGCCACCAC ACCTGGCCTG GAAGGAACCT CTTAAAATCA GTTTACGTCT TGTATTTTGT 60

TCTGTGATGG AGGACACTGG AGAGAGTTGC TATTCCAGTC AATCATGTCG AGTCACTGGA 120

CTCTGAAAAT CCTATTGGTT CCTTTATTTT ATTTGAGTTT AGAGTTCCCT TCTGGGTTTG 180

TATTATGTCT GGCAAATGAC CTGGGTTATC ACTTTTCCTC CAGGGTTAGA TCATAGATCT 240

TGGAAACTCC TTAGAGAGCA TTTTGCTCCT ACCAAGGATC AGATACTGGA GCCCCACATA 300

ATAGATTTCA TTTCACTCTA GCCTACATAG AGCTTTCTGT TGCTGTCTCT TGCCATGCAC 360

TTGTGCGGTG ATTACACACT TGACAGTACC AGGAGACAAA TGACTTACAG ATCCCCCGAC 420

ATGCCTCTTC CCCTTGGCAA GCTCAGTTGC CCTGATAGTA GCATGTTTCT GTTTCTGATG 480

TACCTTTTTT CTCTTCTTCT TTGCATCAGC CAATTCCCAG AATTTCCCCA GGCAATTTGT 540

AGAGGACCTT TTTGGGGTCC TATATGAGCC ATGTCCTCAA AGCTTTTAAA CCTCCTTGCT 600

CTCCTACAAT ATTCAGTACA TGACCACTGT CATCCTAGAA GGCTTCTGAA AAGAGGGGCA 660

AGAGCCACTC TGCGCCACAA AGGTTGGGGT CCATCTTCTC TCCGAGGTTG TGAAAGTTTT 720

CAAATTGTAC TAATAGGSTG GGGCCCTGAC TTGGCTGTGG GCTTTGGGAG GGGTAAGCTG 780

CTTTCTAGAT CTCTCCCAGT GAGGCATGGA GGTGTTTCTG AATTTTGTCT ACCTCACAGG 840

GATGTTGTGA GGCTTGAAAA GGTCAAAAAA TGATGGCCCC TTGAGCTCTT TGTAAGAAAG 900

GTAGATGAAA TATCGGATGT AATCTGAAAA AAAGATAAAA TGTGACTTCC CCTGCTCTGT 960

GCAGCAGTCG GGCTGGATGC TCTGTGGCCT TTCTTGGGTC CTCATGCCAC CCCACAGCTC 1020

CCAGGAACCT TGAAGCCAAT CTGGGGGACT TTCAGATGTT TGACAAAGAG GTACCAGGCA 1080

AACTTCCTGC TACACATGCC CTGAATGAAT TGCTAAATTT CAAAGGAAAT GGACCCTGCT 1140

TTTAAGGATG TACAAAAGTA TGTCTGCATC GATGTCTGTA CTGTAAATTT CTAATTTATC 1200

ACTGTACAAA GAAAACCCCT TGCTATTTAA TTTTGTATTA AAGGAAAATA AAGTTTTGTT 1260

TGTTAAAAAA AAAA 1274

(2) INFORMATION FOR SEQ ID NO: 78:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1133 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:

AGGATTTTTC CTTGTTCAAC CAAAATCTGA GCATTCTTTC TATGTTGAAA ACACTGAAAA 60

ACTAATTTWA GTTAATGAAC TAGAAAGAAT ATTGATTTTW AAGAAACAGA AAAATACTAC 120

TTATTTTCCT TCTCAAATAA CGTTTCTTTC AAAAACTTCT GGCTGAAGTA TAACATGCTG 180

GTAGTTAACA TAAATCTTGT CTTTCTCTTG TTCTTTATCT TTCTTTGTTA TTTAGATGCT 240

TGTATAAATG TCTTTTGTTT TTATTAAGTG CCTAATTGAC AGAGCTTAAT TTGAAGAAGT 300

GCCCTAATTT ATTGACCACT TAAGAATTGC CTTTATTGGG GTATTTTATT TGTTCCTGCG 360

TCTTTTTGAT GTTGTTCAGT CTACTCATCC CTGTGAGTAT GTGTGGGGGA CAGCTGATAG 420

AAGGGAGGAG AGTGTGTCTA TGCTCAGGAT TGCCCTTTAG CCACTCAGCC AGAGATCCAC 480

AGGGAGCAAC AAGGACAGTT TCACATGCTT AGACTTTCTT GGAAGAAACA GTGAGGAGGA 540

GTAAGTCGTG AGTAGTGTCA AGCTGGATGT AGAATTGTCC TAAGGCAGTT GACCCCACCT 600

TCCAACATGT TTTCACTTTA TTTGCCCCTC CCTACATTTG GGTTAGGTTC CATTTGGATT 660

TGCAGCAATA ATGACTTTAT TTCTCTCTTG GTCAGGATTT GGCACATAAA ATCCTTTTAT 720

TATAGAACTA GCTATTTTAG TTACATAGTA ATGTAACTAA TGGAGAGATT TATAGAGAAT 780

TTTGKTTTTG CTGTCATATA TGTCCATTTT GGAGACAGAT ATGATAGAAC TAGAAATTAA 840

GTTGCATTTC TGCAAGTGCC ATTTGAATGA ACTTCAAGTA TCTTCTTAAT TATTAAATTT 900

TCTGATGAAG GCATTGTAAC AAATATATAG TATTATTAAA TCTAATTAAT ATTTGGAAAT 960

ATTAATAAAT AGGTATTTTA TTTACTGTAA AAAGTCAAAC TTCATTATGT AGATAAATCT 1020

TATTCTTTTC ATTCTTTCCC CTGTTTACAT CCTTTTTACA AAGCTTAGTC ACCAATTAAA 1080

GCTTTCCTAT CAAAAAAAAA AAAAAAAAAA ACTCGAGACT AGTTCTCTCT CCT 1133

(2) INFORMATION FOR SEQ ID NO: 79:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 661 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

GAATTCGGCA CGAGGGGAAA AGGATGCTGA ACGAGAGCAG AAAGCCTCTT TCCTTTGCTT 60

CACGCCTTTC CAGTCTTTAT TTTAAACTCG GGTTCCCTTT CTGTGGTCGC AGCAACCTTT 120

ACTCCACCTG CACTGCTGCT CCTGGGGGCT CCCCAGGCCT CCCTCTGCCT TTCTACCCAG 180

TGGCTGACGG GATGCCTGTC TTGCCTGGAC GCACCACTGC TCTCCTGTCC CTCACCTTGG 240

CTTTTGCTGT GCCCTGCTCT GGGGTTGAAG CTGGCCCATG TGTCCCCCGG AGTCATGGCT 300

GCTCCTCCTG GGAGGCCTCT GTGTGCGTCA CGTCTTCCAC ACCTGGGGGC AGCTGGCGAG 360

CCCGTGCTCT GTTCCCCTCG GCTGCTTGGC ACAGAGYTGC AGCCTGGGAY TCTCCGTGGA 420

CCCAGACTGG GGATTTTGCC AGGGGGGCGA TGGGAGGAGC AGGTGCTTTG CCTGGCGGCT 480

GTGTCTGCAT TTCTGGACGC CCCAGAGCAC AGAAGTTGCC GGCACTTTGA GGTCTTCCTC 540

GGCATGTGCC AGATTACATG AGTGACGGCT GGGAATATGT TTTCTTTTTT GTAATGGAGG 600

CGTGTTTCAC ATATAGTAAA GCTCACCAAA AAGTAAAAAA AAAAAAAAAA AAAAAACTCG 660

A 661

(2) INFORMATION FOR SEQ ID NO: 80:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1378 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:

ATTGGGTACC GGGCCCCCCC TCGAAGTTTT TTTTTTTTTT TTTTAATGAA AGCTCTCAAA 60

TAAGCGATTT TATTCCTATC CATGATTGCA GACATTTACA AAACCATAAC ATCTGAGTTC 120

ACCTTAAAAA ATAACTTATA TAAAGCAGTG ATATACACAG CACAAAATAG TTCAGGGAGG 180

GGGCAGGAGC AACTTGTAAT AATTAAAATG TAAACGTGAA AAAAAGGATG GAATAAAAGT 240

CCCTACTTAT TTCTACTTAA GATGTCATGT GATAATATTT TACAATGTCC TGTGGGTCAA 300

TGTATGTATG TGTATATGTC TGTATAACAT ACACATATAC AGTACATTCT CTTTCCCACA 360

CATATACATA CACACATAAT TATTTGCAGT TCAGTTTAGG GCAATTCTAA TATGCCACTC 420

CGTACAGTTG TTTGAATCAC ATTTGGACCC GCTTTCTTCA CAAAAGAGGG GAGAGAGCAG 480

GAAATAAAAA GGTTGGTTTG GTGTGACTGA GATTCCTTTG TTTAACTGTA CACTGTGATG 540

AATAATTTTC TTCCGTAGTA GTTCTGTGAA GGGCTGACTC ACTGTGGTTT TCATGAGGAG 600

ACTTGGTAAT GGATCACACG CTCATTGTCA TGCTAGGGGA GTAACTCTCA CTCTGAAAAG 660

GATTTAAGAA ATTTCCCCCC ATTTCGCCAT CATCCCTTGG AGTGCCCGGT TGATTACTCA 720

GGCTCATATT ATTGGGAGAA TTCTTGGAAA TACTGTCCAT ATCTCCTGAG CCTAAAGAGC 780

CATTCATGTG ATGTGACTCC ATTCCTCCTA ATCCACCCAT GGGACCATCT GACCCAGGRC 840

CCATTGGAAA ATTAGGTCTG TTAGGTCCAG GAGGTACTGC ATTCATTAAA GTATACATGT 900

TATCACCAGA GTTGGTTGAA TCTGCTGGAC TAGGCATGAT GGGTGTTCCT GGTGGCCCTC 960

CACCTCCTGG AGGACCTACA TAATTCCCAG GAGATGCTGA GGAGTATGGT ATTGAATTGG 1020

CATTTGTTGG GTTTGGCCAA GGTCTACCAC CACCTGGACC CATGTTCATT CCAGGCATTC 1080

CAGGGCCACC TAAAGCATTC AGTGGGGGTC TCATTGCACC TCCATAGTTC TGTGGTCCTA 1140

AGGGCACCAT TCCTCTTGGA GGAGTCATTC TCTGCATTGG CCCACCCATA TTTGGATGTC 1200

CTTGTTGTCG AGTTGGATCC ATTCCACTGG GGAGTAATGG CTGACTTCCT GGGACACCTC 1260

CAAGTGCCTG ATTAGGTATC CTCAATGGGG GCCTTGGACC TCCAGGGTAC CGAGGTGACA 1320

TAAAAGGGTA ATCATGGAAG GCTTTTGCTT CACTTGAGTG TTCACATGTT TCACGTCT 1378

(2) INFORMATION FOR SEQ ID NO: 81:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1440 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 81:

ACTTTGTCCA AATGTGTCTG TCACATGTAG TCAGCTGNAG NAATTTAAAA TGAATTGCCA 60

AGTGAAGAGT CTGTGGATTA ATTGGCCGTT AATTAACAGG CTTTATCAAT GTGTCCTCAA 120

GGGAGAGGCC CAACCCTAAT TAAGGAGCTA AACTTCCTGA GTGAGGGGCT GTGAGGATGG 180

AGGTGGAGGA GGCATCTGGG GCGGGTGGTG GCCGGGCCAG CAGATGGCGC CTCCCTGGCT 240

GAGCTGCCCG CACCGCCAGT TCCCTCATTT CCACTCAGGA AGGCAGAGAA GGCAGAGTGA 300

TCTCCTCAAG GAAGAGCTTC CCCAGCCTTC GGGAGCAGCT GGCAGGGCGT CCGGGAATAA 360

GCCCTACACG CCGCCGCCTG CCTCCAACTC ACTAACCCTG CGCCTCTTGT CTTTCAGATT 420

CAACGCGTTC AACAGAAGCC ATCCCCAGCC CAGCTTAAAT TATAAAGATA GACAATAACT 480

CTGTTCCAAT CTGCGTGGTG CTTCTTTAGT AAATACTGTA CAGATTTTAC CATGGAGAAC 540

TTTTTTTTTA GTTTTTACCT TTTCTTAATT ACCCTTATTC CGAATGGACG AACACTTTCT 600

ACCACTGCTG ACCATTGTAA AATACCGTGT ATATAAATCC CATTGAAATA ATGCCCTGGA 660

ATAGAACATC TCAAATGCTG CTTAATTACA GACTCAGGTC GATTACTTGT ATTTCATGTA 720

ATGTTCCTCC AAGTTAGACA TCTGGTGCAA GACCAACCGG GAGACCATGG AATTGTCAAA 780

AGTACAAACT GACAGTGTGT ATATTTAATT TAAAGACTTA TTTAAAAACT CACAAGCTCT 840

CACCTAGACT TTGGAGAGCA GTCTGTTTTC TGTAATGTCT GATACTAGAA ACTAATTTGC 900

TTATTTTAGT TGTATTCAAG ATTTGAAGAT GTATTTTATA GACAAGTTCT GTTTTTGAAC 960

TTTGTGGAAC TGTTCCAATC AATCAATTTC CCAGTTATGA TGAGTATTTA CATTATGAAT 1020

GTATAACCCA GACATGATTT GTAAAGCCGA CAGTATGTTT CTATTACACA ACACTTTTTG 1080

ATACAGCGTC TCTTGTCTTC ACTGATACTG GAGTCTCCGT TGTCTGCNNG GTCCCTTCGA 1140

GTTTCTAGTT ACAGACACAA TCATACTGTG ATTTTATTTT TAATATGGAT ATGCTATCAA 1200

ACTGTGATAC ACTTATAATT CACTGGTCCT GCATCAGGAG ATGGAGTGGG GAAAACTGTA 1260

TTTAATACAG TTTGTATCTG AATAATCTGT ATGGTTTATA CAGTTTGTGT TGTTCAGAGA 1320

TGTTTAAAGT TTGATCTTTG TTTTTCTAAA GATTAAAAAA GCACTTGCCC CACTGTAAAT 1380

ATACAGCATG TAAAATTTCT RTAGTATATA AATGGCAGCA AATCACAAAA AAAAAAAAAN 1440

(2) INFORMATION FOR SEQ ID NO: 82:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1381 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 82:

CCCGGGCTGC AGGAATTCGK YACGAGGCCA GCAGTTGCTC CCAGTTCAGG AGGTGCTCCT 60

GTACCCTGGC CACAGCCCAA TCCTGCCACT GCTGACATCT GGGGAGACTT TACCAAATCT 120

ACAGGATCAA CTTCCAGCCA GACCCAGCCA GGCACAGGCT GGGTCCAGTT CTGACCTGAG 180

CACGGTTTTT CCTCATGTGA CTTCTGGGAA GGCGCTCCCT CATCTGGGCC AAAGGAAGGA 240

GGACGAAGCC CTCCTCAGCT GGCCTGTGTT TGGGGCATGA ATCTCTCCTC TCCTCCTTGT 300

CTGGCTCTGT TGACAAACCG GGCATGTTTG GCAGTAAATT GGCACCGTGT CACACTGTTT 360

CCTGGGATTC AAGTATGCAA CCAGAACACA GGAGAAGAAA AGCTCCAGGA TCCCTGTCCC 420

CATCTGTCCT CTTGATGTGA GAGAGACTCT GAGACTTCTT CCATCGCAAT GACCTGTATT 480

AAACACAAGC CCCCCAAGCA AAAGAAGAGG TTGAGTTTGC TGCCAGGATT CAGATCAGCC 540

CTTCCCAGGG TCTGCAGGTG TCACATGATC ACAGTTCAGC GGGAGGCTTT CCGTACCCAC 600

ACTGGCTGTA GCACTTCAGT CCATCTGCCC TCCAGAGGAG GGTTTCTTCC TGATTTTTAG 660

CAGGTTTAGA GGCTGCAGCT TGAGCTACAA TCAGGAGGGA AATTGGAAGG ATTAGCAGCT 720

TTTAAAAATG TTTAAATATT TTGCTTTGCT AATGTGCTGA TCCGCACTAA CTCATCTTTG 780

CAAAAGGAAC TGCTCCCTCG GCGTGCCCCA GCTGGGGCCT CTGAAGGGAT TCCTCACTGT 840

GGGCAGCTGC CCTGAGCTTC AGGCAGCAGT GTTCATCTCT GGCCAGTTGT CTGGTTTCCA 900

TGTATTCTAG GCCAGGTAGG CAACACAGAG CCAAGGCGGG TGCTGGAAGC CAGACGGAAC 960

AGTGTTGGGG CAGGAAGGTG GATGCTGTTG TCATGGAGCT GTGGGAGTTG GCACTCTGTC 1020

TGCTGGTGGC CCTCTCGGCT CACATGTTCA CAGTGCAGCT CCTGGCAGAC TTGGGTTTTC 1080

TCTTTGGTGG TTTCTAAAGT GCCTTATCTG CAAACAACTT CTTTTCTCCT TCAGGAACTG 1140

TGAATGGCTA GAAGAAGGAG CTCAGTAAAC TAGAAGTCCA GGGTTGCTTG GTTTACTGGT 1200

TTATAAGAAA TCTGAAAGCA CCTCTGACAT TCCTTTTATT AACTCACCTC TCAGTTGAAA 1260

GATTTCTTCT TTGAAAGGTC AAGACCGTGA ACTGAAAAAA GTGTTGGCCT TTTTGCGGGA 1320

CCAGATTTTT AAGATAAAAT AAATATTTTT ACTTCTGTCA AAAAAAAAAA AAAAAAATNT 1380

C 1381

(2) INFORMATION FOR SEQ ID NO: 83:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1706 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

ACTGCACCAC TGCCCAGGTC TCCCGGCTGG ATGAAGACGT GGTCCATGAG GAAGCTGGCT 60

AGCTCAGACT GGAGAGTAGC TTCAGGAAAA AAGACAAGTG GCCTAAGGAA ATCACGGCCC 120

CCAACTATCA TCTGAGGGCT AAAGATGAGA AGTAGATCAC TTAATAAGAC AAAAGCCTGT 180

AGGGGGAAAA GAAAGGATGT TTAAAAGGAC AGAATGTTTC CCAAGGTAGA AATGACACTG 240

TCAATTTCTC CTTGGAATGG GGGCAGGGAT ACTCGCCTTG TTGCTCCCAC TTGAGTCAGT 300

ACTCACCTGC TCCTGGATCT CAGTATCCAC ATCTGAGAGG CAACTCTGGC AGAGTTCACA 360

GAAGGCCACC ATTCTGTCCC TCAAACTCGA CAGCTGCTTC TGTGGGCACA GTGGCTTGAA 420

GGGGAAGAAT GAAGACACAG ACTCCTCTGT TCCCATTATC CCATCTAAGA CCCACACTCA 480

CCTGGGGAAG CATCTGATTT AGAAATGTGG GTTAGTGTCC AGAGAATGGA AAAATAGACA 540

AGAGTCAAGG CTGGCAGGAT AACCTGTAAC AACAAAGGGT TTGAAAAATG AGGTTTGGGT 600

TAGGAGAGGG AGAGACAGAT AGCCAGAAAC ACACCAGTGA AGAGGAGAGA AAATGAGTAA 660

AGGGAGAGCT AATTCCTTTT CCAGTGGAAA ATGAGTGATA TTCTGGACAT TCTTCAGAGG 720

CATCTACACG AAGTAGAAAT GTCACCGCTC CCTAATTTAC TCTACGTCTT CTAGAATCCC 780

TCAATATTAT CCTTGGCTTC CAGGAAATCC AAGAAGACCC TGGAAGTAGA GTCCACCTTC 840

TAAGAGAGGA ATGTAAGAGG TGACCCCCAC CCACCTGATC TTCCTCGCTT TGTCCACTCC 900

ACGCACTGAG ACTTGACACA CCTAGTGGCC ACCTAGAACG TAGGTCCTTA AAATYTAGCC 960

CCCCAGCCCC CAACCCATCT CTAGCCTGTC CACTCACCTG GTGAGGAACY TYTCCTGTGT 1020

CCACAGCYTT CTGCAGGAGT TGGCAACATG GCTCATAGAG CTCCCAGCGA GTCAGGTCAT 1080

GAGTGCTTTG GGGGAGAAAG GGGAATGTTA TACTGGAAAA GAACAGAGGG AACCAACTCC 1140

ACAGACACCA GTAAAAACGG GATGGGGAAG AGGAGGAAAG CCACTCACTT GTAGAAGGCA 1200

GAGAGGCGTT TCAGAGTGGC TGCCAGATTA TATACCTCAT CCTCATCTAG GAAGGACGAC 1260

TGAGAAGGAA AGAAGATCCA CAATAGCATT TCCCCCAGAA CTCATCAGTC CACATCCCCC 1320

GTCTTGCAGC CCCTCCCACC CTTGTTTGGG GTGTCCCATT GTCCAGCCCC AGCTCCTACC 1380

TGTAACAGCT CTTCAAGCTC CTGCTGGAAR CGGTCAGTCA GCAAATCTAC TAGCTGGCTG 1440

CGGGCAAAGT CCGCCCGGCT GAAGAAAGTG AATTCGGGAT TACAGAGCAG GTAAGAGCAT 1500

GCGCCCCAGC CTCAAGCACC GCTGGCTCTG CATGCTTCAC CACCACCTCC TGGAGTTGCT 1560

GCAGGAACAG CTCCAGGTGC TGAGAAGAAA AGGCAGAAGA TGGTGTGCTG TGGGGATGGG 1620

AGGAGGACAC TCTTCTGGCG GGAAGTGGAA CGGGGTTAAA AGCATTAAAC TTCAAGGATA 1680

AGATGCCTAA RAAAAAAAAA AAAAAA 1706

(2) INFORMATION FOR SEQ ID NO: 84:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 573 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:

GAATTCGGCA CGAGCTTGGT AGCCTTAGAA CTGCATGAGC TGCTTTACCA CTGGGAAACA 60

CGAGCACAGC CTAGCTTGAT TTTGTATGTG GTATCAGATC TAAGGTGGAT GGAATTCAGG 120

ACTTCCTGTC TACTCTTTGA TTTTGTTTTA TTTTTAGAAA TGTTTTATTT TGTTTTATTC 180

ATTTATTCAT CTTCAGAGAC ATGGTCTGGC TCTGTTGCCC AGGATGGAGT GCATGGTGTG 240

ATCATAGGCC ACTGCAGTGT TGAGCTCCCG GGCTCAGGCG ATCCTCCTGC CTCAGCTYCC 300

TTAGTAGCTG GGACTATAGG CACATGCCCT ACCATGCCTG GCTTTGTCTA CTTTTTGAAT 360

GATGTCYCAA ACTAGAAGGT CTATTAATTT AAAAAATTAA GGATAGCATG CCATAATTAA 420

AAATAATAAC AGTGGGAAAA GGCACCTTCC AATGATTCAG ACATCAACTT GTGATTTAAA 480

AAAACGAAAA ATAAATAATA GGAAAAAAAG GGGAAAAAGT TAAATAAAAA TAAAATTAAA 540

AAAAAAAAAA AAAAACTCGA GGGGGGCCCG GTA 573

(2) INFORMATION FOR SEQ ID NO: 85:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 684 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:

CTCTTTGGCT GTGTCTACCT CCTTCATCTG CTGCGCCGAC ATAAGCACCG CCCTGCCCCT 60

AGGCTCCAGC CGTCCCGCAC CAGCCCCCAG GCACCGAGAG CACGAGCATG GGCACCAAGC 120

CAGGCCTCCC AGGCTGCTCT YCACGTCCCT TATGCCACTA TCAACACCAG CTGCYGCCCA 180

GCTACTTTGG ACACAGCTCA CCCCCATGGG GGGCCGTCCT GGTGGGCGTC ACTCCCCACC 240

CACGCTGCAC ACCGGCCCCA GGGCCCTGCC GCCTGGGCCT CCACACCCAT CCCTGCACGT 300

GGCAGCTTTG TCTCTGTTGA GAATGGACTC TACGCTCAGG CAGGGGAGAR GCCTCCTCAC 360

ACTGGTCCCG GCCTCACTCT TTTCCCTGAC CCTCGGGGGC CCAGGGCCAT GGAAGGACCC 420

TTAGGAGTTC GATGAGAGAG ACCATGAGGC CACTGGGCTT TCCCCCTCCC AGGCCTCCTG 480

GGTGTCATCC CCTTACTTTA ATTCTTGGGC CTCCAATAAG TGTCCCATAG GTGTCTGGCC 540

AGGCCCACCT GCTGCGGATG TGGTCTGTGT GCGTGTGTGG GCACAGGTGT GAGTGTGTGA 600

GTGACAGTTA CCCCATTTCA GTCATTTCCT GCTGCAACTA AGTCAGCAAC ACAGTTTCTC 660

TGAAAAAAAA AAAAAAAAAA AAAC 684

(2) INFORMATION FOR SEQ ID NO: 86:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1036 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86:

TGGAGGCAGA TGCACAGGAG AAAGGTTCCC GTCCGCACCC TCTCAGACCT GAGGCTGAGC 60

TTGCAGTGAG GGCTTCTCCT CGGCCCCTCG CCCGCCCCCA GAGCTGCCAT CCCTGCTGTT 120

ACAAGCCAGA GGAGCCCGGA TGTGAGGCCC CAGATCACCT CCAGGGACTT GGGGTTCCCA 180

TCTGAAATCC TTTATTTTTG TACCATGGGG TGGGCCCCGG GCTGAGAAGG AAGAAGCACC 240

CTCTCCCCGG CCTCCTCTGT CTGCACCCGT GGGGCTGTGA CTTACTCCTG CCTCCAGGGG 300

CGGGGCGGGG CCCCCTGGGA CCTCTTAAGG CCCAAGGTGG GCCCCAGGAC CTYTGGGCAG 360

AGTGGAYTGC TCATGGCAGA TGTGTGGCAA TGTCTGGCTG WGTCTTTCCG GCAMCTGCGT 420

YCCCTYTCCC GGGYTCCCCT GCTGCATGGT GGATGTGCTC CTTCCTGGCC CGGTCACATT 480

GCCTCCTTGA GCCTTAGTCC AGGGGGTCAC TYCTCCCACC CCACCTACCT CACAGGGTTG 540

TTGTGAGGGT GCACAGAGGA GCAAAGTCCC TGAAGGCCCT CAGGCAGTAT ATAGGGGCCG 600

CCCACCTTCA GCTGCCCTGG GATGGGAAGG ACCCAGCCCG ACCCCTGGGC ATAACACTGT 660

GTTTGCAAAT GGAGATTCAG GTATTGGGGA TGCAGGTTGT GGGGAGCTGG CCTGGCAGAG 720

TAGGGGTAGT TGGCTTGGCC TTCTCTTTGG TGATCCCACC CCCAGCCATT TGCATTGCTG 780

GCCCAGCGCC TGGCCTGGGG GGCGGGGAGA GGCAGCAGAA GGGGCTGGGC AGGGGCGGTG 840

GAGGACTCAG GAACTGCCCG GGGAGAGTGG GTATGGCGGC TGAGCCAGGG GCCCTCCTGT 900

GTTTGACTTC CCGGGATGGG TCCTTGCTTC TCAGCTGTGT CCGACCCCAC CATGTAATAA 960

AACCCAAAGG AACAGCAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAN 1020

CCCNGGGGGG GNCCCG 1036

(2) INFORMATION FOR SEQ ID NO: 87:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 908 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 87:

TTAAACAAAT GGAATCATGC AATATGTGAC CTTTTGCGTC TGGCTTATTT TATTTAGCAT 60

AATGTTTTTG AGGTTCATCC AAGCTGTAGC ATGTATCAGC ACCTCATTTC TTTTTCTGGC 120

TGAATATTAT TCCATTATAT GGATTTACCA CAATTCATTT ACCTATTCAT CTTTTGTTTC 180

TGCTGTCTGG CTATTGTGAA TAATGCTTCG ATAAACATTC ATATACAAGT TTCTATGTGG 240

CTTTATGTTT TCATTTCTCT TGGCTATCTA CATGGGAGTA GAATTCTAGG TCATAATATA 300

ATTTTATGTT TAACTTCTCA AAGAATTGCC AAAAGGTTTT TCATAGTGGC TGCATCATTT 360

ACATTCCCAC CGGCAATGTA CAAGGATTTC TATTTTTCCA TATCCTTGCA CTTACCAACA 420

CTTCTTTTTK GTWATWATTT TGTTTTTTCA TTATTGCCAC CCTAGTGGAT GTGAAATGGC 480

ATCTTATTGT TTTGATTTGC ATTTCTCTAA TGACAAATGA TATCATACTT TTTTTATGTG 540

CTTACGGATC AAAGGTATTT CCTTGGAGAA ATGTCCCTTC AAGTCCTTTG CCATTTCAAA 600

ATTTGGTTAT TTGTCTTTTA TTATTCAGTT TTAAGAAATT CTGGCCAGGC GCAGTGGCTC 660

ACCTGTAATC MTAGCACTTT GGGAGGCCAA GGCGGGCAGA TCACTTGAGK TCAGGACTTC 720

GAGACCAGCC TGGCCAACAT GGTGAAACCC CATCTTACTA AAAATACAAA AATTAGCTGG 780

GCGTGGTGGC AGGTGCATGT AATCNTATCT ACTCAGGAGG CTGAGGCAGG AGAATCGCTT 840

GAACCCAGGA GGCGGAGGCT GCAGTGAGCC AAGATCACGC CATTGCACTC TAGCCTGGGT 900

GACACAGA 908

(2) INFORMATION FOR SEQ ID NO: 88:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 655 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:

TGCACTGGTT CCTTCTCCCC AGCAAATACT GCCTTCTTGT TTTTCTCTGA TGTGGCAGGT 60

GACTACAAAA TCCGCCTTGG TATTCTTCAA ATGCATATAT ATTCCTTTCT TGTCAGCTCC 120

CTCTCTTCCT AGATTAGAAA ACTGCCTCAT TTTCTGCTCA CTGGATGTGC AGTCCCAGCT 180

TGTCTTCCTC TCCTCCCCCC CTGTTGCAGG TGTTCTTTTT TTTTTTCTTC TCTCCCCACT 240

GGGCAGCAAA AGTTGTTCCA CAGTGGAAAW TTAGGCATCC TCAAGTTTCY TCCCAGCTTC 300

TGCTGTGTTT TCTTAGAGTA AATTGCCAAT TTCTGTTTTT ACAGGAAATC CTTTTTTAAA 360

AATGGAATCA GTGTGGTCCC CATCTACTCT GCAAAAATTG CATTTTTCTC TATTTTCAAA 420

TGAGATTTGT TCAAGTTTCA AAACCACGTG AAATAATAAA TGTATAGTAG TTTTCTTTTC 480

CTTGGGCATT GCTWGATATG TGAAATGGGT TTATGAAAAA TAATAAAATC ATAACGCTAT 540

TTGTTTGACT TTCAATTTCA TGGGAATTTT TCTCAGCTAA ACTCTAAATG GTGATTARGC 600

AAAAAAAAAA AAAAAAAACY GRAGGGGGGC CCGGTACCAA TTCGCCCTAT AATGA 655

(2) INFORMATION FOR SEQ ID NO: 89:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1102 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 89:

TTTTTTTTTT ACCATTTAAA ATAAAATGAA AGTGACCTTC TGTTTATAAA AATCTTTGTC 60

TGCATCTCTG CTTATTTCCT TAGAAGAGAT TCCAAGAAGC GGTGAGTGAT TTCACGGCAG 120

CAGAGGGTTG GGACATATTA CGGGCGCGGA TCCCTCTTGG AGTGAGATGA CTCTCCGGAG 180

AGATTTAGTC GTCACCCTCG CGTGTGAGGC TGCGTCACAC CCCAGGGATG TGTCTATCAA 240

GATGGAAGAT CTTTTACACG CTCTTGATTT TGTTTGSCTY TTTTTCTATT ACTAGTGAGA 300

AKGAAACTTT TTATATGATT ATTATCCATC ATAATCCAAC ACAAATTACT GCTTCATGTT 360

CTTTTACTTT CCTGTGAAGG TTTTAGTGCC TTTTAAAAAT TGCTATATAT TAAGCTTGTT 420

AATACTTCCA TGCTGTATTT GTGGSCATCA RTTTCCCCGG GNACAGGCNT GCACATTTTG 480

CCTTCACACG CTGGGTGGTT TTTCATTTTC AMTTCTATTT CTCGTTCTTC TATCGTTTTA 540

TGTTCAGACG GGTTTCTCCG TGTAGAAAGC AGTTTATGAA GATTTACTTT CGACAGTCTT 600

CTCTCTACTT TCTACAGTGA ATTCTCTGAT GTGTCTGGGA GTTTGGGGGT CTGGGTAAGA 660

RTCCTCCTCT CACCCTATTC TCTATTACGA TCCACAGCCT CATGCTTTAT GARATTGGTG 720

GCCGGGARCG GGGGAGATTT GCGGATCCCC CAAGCCAGAC TTTATCCCCC TATCCCTGCC 780

TCTGGATCCC ACGTACAGGC CTGGGAACTC CCTGTGGGTA GGGGCCAATG GTCTCGCACT 840

CTCACCTGTA CCCCAGGGCT GGCACAGGAT GGTCAAGGAG AGAGGCTGCC CAAGCGCATC 900

CYTCTGGTGT CCCCCTGACA CGCCTCCAAA GTGAGCAGGT AGGTTTCAAC AGCCCCACGT 960

TGCAGGTGGG AGATGAAGCT CAGGGTGGAG ACCAGTATCT CACAGTTCTC TTTGCATGGC 1020

CGGGTACTTG TTAGTCAACT GATCAAGTGA AAATTCTAGC CCCAGAGGCA GGAGAATCCG 1080

GAACAAAATT AAACCAGCCA GG 1102

(2) INFORMATION FOR SEQ ID NO: 90:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1533 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90:

GGCACGAGCC GNCACGGGCA GCGCCCCATA GCGCCAGGGA CCCCCTGGCA GCGGGAGCCG 60

CGGGTCGAGG TTATGGATCC AGCGGGCGGC CCCCGGGGCG TGCTCCCGCG GCCCTGCCGG 120

TGNCTGGTGC TGCTGAACCC GCGCGGCGGC AAGGGCAAGG CCTTGCAGCT CTTCCGGAGT 180

CACGTGCAGC CCCTTTTGGC TGAGGCTGAA ATCTCCTTCA CGCTGATGCT CACTGAGCGG 240

CGGAACCACG CGCGGGARCT GGTGCGGTCG GAGGAGCTGG GCCGCTGGRA CGCTCTGGTG 300

GTCATGTYTG GAGACGGGCT GATGCACGAG GTGGTGAACG GGCTTCATGG AGCGGCCTGA 360

CTGGGAGACC GCCATCCAGA AGCCCCTGTG TAGCCTCCCA GCAGGCTCTG GCAACGCSCT 420

GGCAGCTTCC TTRAACCATT ATGCTGGCTA TRAGCAGGTC ACCAATGAAG ACCTCCTGAC 480

CAACTGCACG CTATTGCTGT GCCGCCGGCT GCTGTCACCC ATGAACCTGC TGTCTCTGCA 540

CACGGCTTCG GGGCTGCGCC TCTTCTCTGT GCTCAGCCTG GCCTGGGGCT TCATTGCTGA 600

TGTGGACCTA GAGAGTGAGA AGTATCGGCG TCTGGGGGAG ATGCGCTTCA CTCTGGGCAC 660

CTTCCTGCGT CTGGCAGCCC TGCGCACCTA CCGCGGCCGA CTGGCCTACC TCCCTGTAGG 720

AAGAGTGGGT TCCAAGACAC CTGCCTCCCC CGTTGTGGTC CAGCAGGGCC CGGTAGATGC 780

ACACCTTGTG CCACTGGAGG AGCCAGTGCC CTCTCACTGG ACAGTGGTGC CCGACGAGGA 840

CTTTGTGCTA GTCCTGGCAC TGCTGCACTC GCACCTGGGC AGTGAGATGT TTGCTGCACC 900

CATGGGCCGC TGTGCAGCTG GCGTCATGCA TCTGTTCTAC GTGCGGGCGG GAGTGTCTCG 960

TGCCATGCTG CTGCGCCTCT TCCTGGCCAT GGAGAAGGGC AGGCATATGG AGTATGAATG 1020

CCCCTACTTG GTATATGTGC CCGTGGTCGC CTTCCGCTTG GAGCCCAAGG ATGGGAAAGG 1080

TGTGTTTGCA GTGGATGGGG AATTGATGGT TAGCGAGGCC GTGCAGGGCC AGGTGCACCC 1140

AAACTACTTC TGGATGGTCA GCGGTTGCGT GGAGCCCCCG CCCAGCTGGA AGCCCCAGCA 1200

GATGCCACCG CCAGAAGAGC CCTTATGACC CCTGGGCCGC GCTGTGCCTT AGTGTCTACT 1260

TGCAGGACCC TTCCTCCTTC CCTAGGGCTG CAGGGCCTGT CCACAGCTCC TGTGGGGGTG 1320

GAGGAGACTC CTCTGGAGAA GGGTGAGAAG GTGGAGGCTA TGCTTTGGGG GGACAGGCCA 1380

GAATGAAGTC CTGGGTCAGG AGCCCAGCTG GCTGGGCCCA GCTGCCTATG TAAGGCCTTC 1440

TAGTTTGTTC TGAGACCCCC ACCCCACGAA CCAAATCCAA ATAAAGTGAC ATTCCCAAAA 1500

AAAAAAAAAA AAAAAAAAAA ANCCCGNGGG GGG 1533

(2) INFORMATION FOR SEQ ID NO: 91:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 575 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

ATCCTCTGGA ATCTAGGTGG AAGCCACCAA GCCTTCTTCA CACTTGCGTT CTGAGCATCT 60

GCAGACTTAA CCCCATGTGG CAATCACCAA GGCTTATGGC TTGTGTCCTC CAGAACTGTG 120

GCCAGAGCTG TACCTGGGCC CCTTTGAGCT GAGGCTGAAG CCAGAGTCTG AAGCTCAGCA 180

GGGCAGTARG GCCCTGGGCC TGGCCCCTGA AACCATTCTT TTCTCCTAAG CCTCTGGGCC 240

TTTGATGGGA RGGGCTGTCC TCAAGATTTT TGAAATGCCT TTGGAGGGTT TTTGCCTTGT 300

CTTGGATATT GGCTTCCTTT TAGTTATGCT CATCTCTCTA GCAAGTGAAT GTTTCACAAC 360

CTGCTTGGAT TCTTTCTCTA CCACAGARCC AGGCTGCAAA TTTTACAAAC TTTTACACTC 420

TGTTTCCCTT TTAAATATAA ATTTCAATGT TAAGTCACTT CTTTGCTCCC ATATCTGATT 480

TAGGTTGCTG GAAGTAGCCA AGTCACCTCT TGAATGCTTT GCTGCTTAGA AATTTCCTCT 540

ACTAGGTAGC CTGGGTCATC ACACTTAAGT TCAAA 575

(2) INFORMATION FOR SEQ ID NO: 92:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 639 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 92:

TCCTTTCATC TTAAGCACCA CCCGACAGGG CAGGTACTAT TACCATCTCC GTTTGACAGA 60

TNAGGAACCT GGCACAGGAA GCATTTAAGT GGATTCCCCA GGATCGCCCC ACTGTCAGGA 120

GCAGANTCAG AATGGGCCTC AGCATCAGGC TCCCAATCCT GGCTTCTAAC TGCTGCGCTC 180

TGCCCTTCYC TCWCCCCACC TCCCCACTCC AGTGCCTTTG GTCATGCCAC TGCAGCTTTC 240

AGGCCAATAC TGGATTAGCC TCTTAGTGTT CTTGTCCCTG CAGCCATTTC CCCAGGCAGC 300

AATTCCATGT GCCCTCACTG ATGTAGGTGG CTCTTGTGTC ATTTGTCACA TCCTATTGAA 360

TTGTTTATGC ATCTTGTTCA CACTCACAGC ACCCTCCCTC TCACACGTCC TCCTTATAAA 420

AATGTCCCTC AGTGTCTGCT ATGAGCCAGG TGCAGACTTA AGTGACAGGG CTGCTACGGG 480

AAATAAAAAA TTAACAAGGA GCACCTGCCT CTTAATGCAC AGTAACAAAC TATGTTAAGT 540

GTCAGGAAGG AAAGGTTAAG GATGCCAGGA AGGCTTTTAA TAAATAACCT GACTTAGATG 600

GGCAGGTGGT GCTGARGATT AAGAACGTGT TCTTCTCGA 639

(2) INFORMATION FOR SEQ ID NO: 93:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 744 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93:

GAATTCGGCA CGAGAGTGGC TGGAGTCTGG CTGCAGAGGG AAGACATCAG CAGGGAGGGA 60

GCCAGGGCCT GTCACATCTT TCCTCTGGCC ATTGTCCTGG TCTTTGTAAG CCCAGAATCT 120

CCCCTTCCCT GAAGGGAGGC CAGCACCCCA GGAGGGCAGC AGGTGTGCTG TGAGGGTTGG 180

AGTAGTGTGA GAGGTCAGGG TACACTAGAA TGGCCATGGA CACCATGTGG GGGTGCTCTG 240

GGCTGGGCCA CAGAACAGTG TCCTTCCTGC TGCTCCTCCC CTGCAGCTTC CCCCGACCTT 300

GTNGTTTATT TGGTTTGATA CCAATCAGCA GACCCTGCAA GGTGGAAGCT CCCAGGCTCT 360

CAGTCCCACS ACTCTCATGT GCCAGTCACC CNTACTGTAA CTGCCCAATG AGTACTTCTT 420

GCCCACTGCC AAGATAGAGC CAGTTTACCA AGACAGGGGA ATTGCAGTAG AGAAAGAGTT 480

GAATATACAT AGAGCCAGCT AAATGGGAGA GTGGAGTTTT CTTATTACTT AAATCAGCCT 540

CCCYTAAAAT TCAGAGGTGA GAATTTTTCA AGGACAGTTT GGTGGSCAGG CCTAGGGAAT 600

GGATGCTGCT GATTGGCTAG GGATGCAATC ATAGGGGTGT AGAAAAGTWC CTTGTGCACT 660

GAGTCCACTT TTGGTGAGAG CTACCAAGGA GCTGCTGGTC TGCTGGTCCC GGTAGAGCCA 720

TCTGGTGTCA GGAATGCAAA AGTG 744

(2) INFORMATION FOR SEQ ID NO: 94:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 526 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:

GCAGGGGAAT TCGGCCACGG AGGGGTTTCA ACAGGGCCCG TGGGGTGAGG TGCARACACA 60

AAGCCCATAA GTGCTGGCCT GTTGGGACAA ATGAGAGAAA TCCCATAGGG TGGTGATGAC 120

AGCGCAYTCA GCCATCYTAY TCCTGGGGAA AATGAAACTT GTGCTCCTAT CAAATGCTCA 180

GTTGTAAAAC TGGAAAAAAA TTTTAGAAGA CATCTTGTCC AGCATCTGTG TTTATGTCTA 240

TAAAATGTAG AAAACTAAAG CACAGAGATG TTAAATGTTT TGTCCAAGGT CCAACAGCTG 300

GTTAGCARGC TTGGTCTGGT GACCTTTCTA CTGAACCACA GTGCCGCTGG GGGAAGTCCT 360

CAGCACAGAT GGCTGCTGCT ATAGCTGGGG TATGGGCAGT ATTAGTAGTT AACCAGTCAA 420

CCCAAGTTCC CATAGTCTAG GTTCTGCTTC AGCTGGAGGT TAGGGAAAAA CACAAGAAAA 480

TCCCTTACCA CTCTACCAGT GCTGGGGGAT GTACTAAGAG ATCCCC 526

(2) INFORMATION FOR SEQ ID NO: 95:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 426 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95:

GGCACAGGGC AGGAGAGACT TGGTCCATGG GGAGAAGCCT GCAGTATAGA TGGGACCTCC 60

AGGAGCCCAA GTAGCATAGA CCCTGCTGAT CCGGGGCCAT TGAGCCAGAG GATTTGGGCT 120

GAATGTCCCC AGAGACAAAA GGGAAAGGTA GATCCTTTCC CTTAAAGATG AAAGCCATCG 180

CCCGGGCTTG CTTATTGCTC TCTCTCCTGG TCCTTCCACA TGTTGTTTCT GAACATTTGT 240

TCTGGCATCA CAATCCCCGT CATCCTGTCA TCTGGCCCTT CCCACCTTTC CACCTTATCT 300

CTTGCAGTGT CTCCGCGTCG ACCTGGCACC TGGGTGAARG CTTGCTCTTG CTGGTGCCCA 360

TAGCCCCCAG TGTATGGTCT TGAMCTCCCC AGCCATATGG ARACCCACCT CAGGAGGGCC 420

CCTCGA 426

(2) INFORMATION FOR SEQ ID NO: 96:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 844 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

GGCACAGCGG CACGAGATAG GAAGCTTGGC AGGGGCAGCT CCCCCAGTGC GCATTGCCCT 60

GTAACTCGAG CGCCTGGGAG TGGGGAGAGG CTTGGAAATG GAGCAGGGTG GTGGACCTCG 120

TCTTCTCCTG CTCATCCCAG GCCTCCTCCA TAACACCTAC CTAGCACGGC CTGGGGACTT 180

CCCAGCCCAA GGAACAACTG AGAATACTGA GTGCCAGGGT AGCCCTAGCC CCATTTCACA 240

CCTGGGCAAA GTGAGGTCAC TGGATTCAAA CACTCAGATT TAAACCTCCT CTGTGTCTGC 300

AGCACCTGTA TATAACTGCC AGCCTCTGCT GCCCCTCTCC AAAAAGTCTC TGCCCTTGTC 360

TTTGGCACCT GTCTCTGTCC TCCCCATTCT CTGCTCCTCC TTTCTCCAAC TCAGANTCAC 420

CCTGTTAGTT CAGCAAATGT TCATCGAGCT CCATAATGTA GCAGGACAGG NCTGTCTAAC 480

AGATTCTGGN CTTGCAAGGG TGAGACAAGT ACTCTCCATC TTTCTCTCAT CTTCACAGAT 540

GGTCTGCTCA ACAACTTTGC ACTGAATTGT AAATAATTGA TACTGCATAA AACATTGATG 600

TTCTTTAAGG GTAGTCCAGC AAGGTGGCAA GTCTTATAAT GATAACTGCT CAAGGATCTC 660

TCAGTGAAGC ATTTGGGGST GCTAGCTCTG CCTATGGGTG AGGTCAGCTA TCTCACGCCA 720

TCTACTTCCA CNTGCCCCCC CATGCCAGGC TCACCCTGAG CTGAGATGCC TGAGCAGGTG 780

GCAGAAAGGA GCCACCTGGT TTATGCTTCG GGACCACAAA CTCCTCTATC CAGANGACAG 840

TTTT 844

(2) INFORMATION FOR SEQ ID NO: 97:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1985 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:

AGCCCTGCTG AAGTACAGGT TCTTCTATCA GTTTCTGTTG GGCAATGAAC GAGCAACAGC 60

AAAGGAGATC AGGGATGAAT ATGTGGAGAC GCTGAGCAAG ATTTACCTGT CTTACTACCG 120

CTCTTACCTG GGGCGGCTCA TGAAGGTGCA GTATGAGGAA GTCGCTGAGA AAGATGATCT 180

AATGGGTGTG GAAGATACAG CAAAGAAAGG ATTCTYCTCA AAGCCATCGC TCCGCAGCAG 240

GAACACCATT TTCACCCTAG GAACCCGCGG CTCTGTCATC TCCCCCACTG AACTTGAGGC 300

CCCCATCCTG GTGCCTCACA CAGCGCAGCG GNAGAGCAGA GGTATCCATT TGAGGCCCTC 360

TTCCGCAGCC AGCACTACGS CCTCCTAGAC AATTCCTGCC GCGAATACCT TTTCATCTGT 420

GAATTTTTTG TTGTGTCTGG CCCAGYTGCA CACGACCTGT TCCATGCTGT CATGGGCCGT 480

ACACTCAGCA TGACCCTGAA ACACCTGGAT TCTTATCTAG CTGACTGCTA CGATGCCATT 540

GCTGTTTTTC TCTGTATCCA CATTGTTCTC CGGTTCCGTA ACATTGCAGC AAAGAGGGAT 600

GTTCCTGCCC TGGACAGGTA CTGGGGAACA GGTGCTTGCC TTGCTATGGC CACGGTTTGA 660

ACTGATCCTG GAGATGAATG TTCAGAGCGT CCGAAGCACT GACCCCCAGC GCCTAGGGGG 720

GTTGGATACT CGGCCCCACT ATATCACACG CCGCTATGCA GAGTTCTCCT CCGCTCTTGT 780

CAGTATCAAC CAGACAATTC CTAATGAACG GACCATGCAA TTGCTGGGAC AGCTGCAGGT 840

GGAGGTGGAG AATTTTGTCC TCCGAGTGGC AGCTGAGTTC TCCTCAAGGA AGGAGCAGCT 900

TGTGTTTCTG ATCAACAACT ATGACATGAT GCTGGGTGTG CTGATGGAGC GGGCTGCAGA 960

TGACAGCAAA GAGGTTGAGA GCTTCCAGCA GCTGCTCAAT GCTCGGACAC AGGAATTCAT 1020

TGAAGAGTTG CTGTCTCCCC CTTTTGGGGG TTTAGTGGCA TTTGTGAAGG AGGCTGAGGC 1080

TTTGATTGAG CGTGGACAGG CTGAGCGACT TCGAGGGGAA GAAGCCCGGG TAACTCAGCT 1140

GATCCGTGGC TTTGGTAGTT CCTGGAAATC ATCAGTGGAA TCTCTGAGTC AGGATGTAAT 1200

GCGGAGTTTC ACCAACTTCA GAAATGGCAC CAGTATCATT CAGGGAGCGC TGACCCAGCT 1260

GATCCAGCTC TATCATCGCT TCCACCGGGT GCTGTCCCAG CCGCAGCTCC GAGCCCTCCC 1320

TGCCCGGGCT GAGCTCATCA ACATTCACCA CCTTATGGTG GAGCTCAAGA AGCATAAGCC 1380

CAACTTCTGA TGTGCCAGAA ACCGCCCTGA GATCTGCCGG TCATCTCCAT GGACTTCTGC 1440

ACCCCATTCC ATACCCTTCT TCACCTGGGG TACCCCTTCC AGTTTTCCCC TTGCTTCCCA 1500

GGCCCTTGAC ATGGCTTACC TGCCTTCACT CCCAGCACCT TGCCCAACAG GATAAGCTGG 1560

ATCCCCTTGG CCTTCTGAAT ATCCCAGTGT CTTCAGGTTT CCCAAGACCA CTTCCCTGTG 1620

GGCTTCCAAA ATGGCCTTTA TCATTTCTCC AGTCTGTCAC CCTCCTTTCC TGCTCCCATA 1680

CACCCAAGGC TTGTTTCTTC CCCTGTAAAA ACCACTGCCT CAATCTCTGG TTCACTCAAC 1740

TAGTCACCAT GTCCTGAGGC ATGAAGCCTC CTCAGCTCTT GGAATTGCTG GCAAGGGGTG 1800

ACTGCCTCTG AGTCATTGTG TTTTTCAAAG TGATTTCTTT TCTGTAGCTT TTTGACCTAA 1860

GATCTCAGCA ATTTGAACAC TAACCTCTCC CCTCCTGGCT CAAGAATTAC TCCGAAGTCA 1920

GTCTGCAGAA AATAAATATT TAGTATGACA TGAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1980

AAAAA 1985

(2) INFORMATION FOR SEQ ID NO: 98:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1416 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:

ATATGAAGGG AAAGAATTTG ATTATGTTTT CTCAATTGAT GTCAATGAAG GTGGACCATC 60

ATATAAATTG CCATATAATA CCAGTGATGA CCCTTGGTTA ACTGCATACA ACTTCTTACA 120

GAAGAATGAT TTGAATCCTA TGTTTCTGGA TCAAGTAGCT AAATTTATTA TTGATAACAC 180

AAAAGGTCAA ATGTTGGGAC TTGGGAATCC CAGCTTTTCA GATCCATTTA CAGGTGGTGG 240

TCGGTATGTT CCGGGCTCTT CGGGATCTTC TAACACACTA CCCACAGCAG ATCCTTTTAC 300

AGGTGCTGGT CGTTATGTAC CAGGTTCTGC AAGTATGGGA ACTACCATGG CCGGAGTTGA 360

TCCATTTACA GGGAATAGTG CCTACCGATC AGCTGCATCT AAAACAATGA ATATTTATTT 420

CCCTAAAAAA GAGGCTGTCA CATTTGACCA AGCAAACCCT ACACAAATAT TAGGTAAACT 480

GAAGGAACTT AATGGAACTG CACCTGAAGA GAAGAAGTTA ACTGAGGATG ACTTGATACT 540

TCTTGAGAAG ATACTGTCTC TAATATGTAA TAGTTCTTCA GAAAAACCCA CAGTCCAGCA 600

ACTTCAGATT TTGTGGAAAG CTATTAACTG TCCTGAAGAT ATTGTCTTTC CTGCACTTGA 660

CATTCTTCGG TTGTCAATTA AACACCCCAG TGTGAATGAG AACTTCTGCA ATGAAAAGGA 720

AGGGGCTCAG TTCAGCAGTC ATCTTATCAA TCTTCTGAAC CCTAAAGGAA AGCCAGCAAA 780

CCAGCTGCTT GCTCTCAGGA CTTTTTGCAA TTGTTTTGTT GGCCAGGCAG GACAAAAACT 840

CATGATGTCC CAGAGGGAAT CACTGATGTC CCATGCAATA GAACTGAAAT CAGGGAGCAA 900

TAAGAACATT CACATTGCTC TGGCTACATT GGCCCTGAAC TATTCTGTTT GTTTTCATAA 960

AGACCATAAC ATTGAAGGGA AAGCCCAATG TTTGTCACTA ATTAGCACAA TCTTGGAAGT 1020

AGTACAAGAC CTAGAAGCCA CTTTTAGACT TCTTGTGGCT CTTGGAACAC TTATCAGTGA 1080

TGATTCAAAT GCTGTACAAT TAGCCAAGTC TTTAGGTGTT GATTCTCAAA TAAAAAAGTA 1140

TTCCTCAGTA TCAGAACCAG CTAAAGTAAG TGAATGCTGT AGATTTATCC TAAATTTGCT 1200

GTAGCAGTGG GGAAGAGGGA CGGATATTTT TAATTGATTA GTGTTTTTTT CCTCACATTT 1260

GACATGACTG ATAACAGATA ATTAAAAAAA GAGAATACGG TGGATTAAGT AAAATTTTAC 1320

ATCTTGTAAA GTGGTGGGGA GGGGAAACAG AAATAAAATT TTTGCACTGC TGAAAAAAAA 1380

AAAAAAAAAA AAAAGGAAAC TCGAGGGGGG GCCCGG 1416

(2) INFORMATION FOR SEQ ID NO: 99:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1935 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99:

NTCTACCCTA ATCAAGATGG GGACATACTT CGCGACCAGG TTCTTCATGA ACATATCCAG 60

AGATTGTCTA AAGTAGTGAC TGCAAATCAC AGAGCTCTTC AGATACCAGA GGTTTATCTT 120

CGAGAAGCAC CATGGCCATC TGCACAATCA GAAATCAGGA CAATAAGTGC TTATAAAACC 180

CCCCGGGACA AAGTGCAGTG CATCCTGAGA ATGTGCTCTA CGATTATGAA CCTCCTGAGC 240

CTGGCCAATG AGGACTCTGT CCCTGGAGCG GATGACTTTG TTCCTGTGTT GGTGTTTGTG 300

TTGATAAAGG CAAATCCACC CTGTTTGCTG TCTACTGTGC AGTATATCAG TAGCTTTTAT 360

GCTAGCTGTC TGTCTGGAGA GGAGTCCTAT TGGTGGATGC AGTTCACAGC AGCAGTAGAA 420

TTCATTAAAA CCATCGATGA CCGAAAGTGA CCAAGACCAA GGCCCACCAA GGCAGCAGAC 480

TGTTAATCAG ACAAACAGAT CTCTGAGAAG GTGCATCAGC TGCTTTGAAG GCTGAAGATT 540

GTTTTGTATG ATACTGCACA GCATCAGGCA TTTTAAAGCA GATCTTTACT AAACAGGTTA 600

ATGAGCTAAC AAGCAGGTTC TCTCGTCTTT GGGCTCTTTC CTTTCTGAGT TGCATATTCT 660

ATTTTCTTGT CCCCAAGTAG AGACTAGTAC TACAAAAAGG GACCACATTT TTCAAGTATT 720

TCTAAGTATA AAAAACAAAA CAAAAATCTC TTAGGAAATG TCTAGACCTC CATTCTTGGA 780

TTCCCTTTCT TTCCTTTTAT TTTAAAAAAG AACAGTACCC CTCTTTTAAG ATGCTGTCTT 840

ACATTAATGA GCATCTAATG GAAAGAAGGT ATGAGTTGCA CTGAGGATTA GAATAGTGGT 900

GCGTTAGTGG CATTATCTAT AAATACACTC ACCTAAATTG AAAGCTAAGA AGGAAATGTA 960

AATATAATAT ATATTTATAT TTGATGTAAT ATGGACATCT GCAGATTCTA ATAAACAAGG 1020

ACTATTGCTG ATAGTAGGCT GTGACATACT GTCTTGTGAA ATGGTTTCCT TGACAAAATT 1080

TAAGCTGAGC TTAAAAGCAA AAAAACAAAA AGTACACAGA AATATTTATT AAAATGTAAT 1140

ACAGTTTATT GAACTTTCTA GGTATGGAGT TTGATGGACA GGGCTGCCTY TAATGAGTGT 1200

GAAGGTCACT AAGTCACTTA GACATCTCAC CGTGGAAGTT TGTGAGCCTG CATTAGGAGA 1260

TAGACTGATT ACCATACATG ACATAAAAAG GAACAGTGGA TAGCTCATAC TTTATGGTGG 1320

TTCTTCTCCT CCGAAATAAT ATACTGCAGA AATCCCAGAC AGAGCTCCTT ACAAACCTTT 1380

AATTGTAATA TATTTTTGAT GATTATTCAC ATTGAATGCA CAGACCAAGA ATTCAGTGAA 1440

TGTCATTTTT TAAAAAACTA ATTTGTATTG TCTGCTCTAG TGATACAAGT TTTACTAGTG 1500

ATAAACTATT TTAATCAACC ATACTATTCT TATGGAAAAA AATATCTATT TTGGCAGGTT 1560

TCTGTGCCTT TATTTCCCTC TTCTGAAAAA AAGTCTGTGT TTTCATAGTT TGGTTTGCAT 1620

TGTATATCAA TAATTAATCA GGAATGGGTT TTGGTGCCTG AAAAATTGGC CATGGAGGCA 1680

CACCAAAGCT TCAAGCACAA GTCTTGTACA TGGGCCATCA CTGTCTGGTT TCACTTCGTG 1740

TGTTTCCTAA ACACATTTAG CTGCTTTTTT AACAAACTCA GCCCCATACT TGAGTCCCTT 1800

GTTGTTGGGA GCATTTCCAG GCATCTTTTA AGGGAACTGT GACAAACAGC CTCGGGCAGA 1860

TGAACACGGA GGCTCTCTGT TGTCTGTCTC TGAGATCTTT GTGTCTGGGA ATGCCTAAAG 1920

NTTTTGNTTT TTTTT 1935

(2) INFORMATION FOR SEQ ID NO: 100:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 599 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:

GAATTCGGCA CGAGCGTCCA CGCAGCCGCC GGCCGGCCAG CACCCAGGGC CCTGCATGCC 60

AGGTCGTTGG AGGTGGCAGC GAGACATGCA CCCGGCCCGG AAGCTCCTCA GCCTCCTCTT 120

CCTCATCCTG ATGGGCACTG AACTCACTCA AGACTCCGCT GCCCCCGACT CCCTGCTGAG 180

AAGTTCAAAG GGCAGCACGA GGGGGTCTTT GGCTGCTATT GTCATCTGGA GGGGGAAGAG 240

TGAGAGCCGG ATAGCCAAGA CCCCAGGCAT TTTCAGAGGT GGCGGGACCT TAGTCCTACC 300

CCCAACACAC ACCCCTGAGT GGCTCATCCT CCCTTTGGGC ATAACGCTGC CCTTGGGGGC 360

TCCAGAAACA GGCGGTGGGG ATTGTGCCGC TGAGACCTGG AAGGGCAGCC AGCGTGCCGG 420

CCAGCTGTGT GCATTGCTGG CTTAATATGC AGGGCTTGGG GGGCTGTGGC CACATGCCCG 480

GCAGGAGGTG AGTGAGGAGC CCTGTGGCGT GCTGGTGTGG GGATCGTGGG CATTTCAAAC 540

GGGCTTGTCG TACCCTGAAC AATGTATCAA TAGAGAAAAA AAAAAAAAAA AAAACTCGA 599

(2) INFORMATION FOR SEQ ID NO: 101:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 784 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101:

GAATTCGGCA CAGAAAAAAA AGAGAGACTG GGTCTTACTG TGTTGCCCAG ACTTGTCTTG 60

AACTCCTGCC TCAGCCTCTC AAGTACTTGG GATTATAGGC CAAGAAGCCA CCATGCCTAG 120

CTTCTTCCTG TCATTGATCC AGACTAATAC TCTGGGGTCA GCCTCATTTC TTCTCTTTCT 180

CACTTTGCAC ATCCACTTGT CACCAAATCK RGTTCATTCT GCATCCTAAG TAAGTCCTTT 240

GATTCCTCCA GTTGTTCATT AGTAATGTCT CAARTGTAAT TTTTTCTAGT AGTTTTCAGC 300

CTGTCTTTCC KGCCTTCAGT CTTAACTTCT CCAGTACATA KGCCACATTG TTGTCAGCAK 360

GATCAWATTT TATTTAAAAA TACTTTACAW AKGTTTATKG CCAAATATTA GRAAATACAG 420

ATTCATGGAA AGAAAAATCA CTGTCCCAAG GAGGTCACTG GCATGGTGAG GTTAAGGGGT 480

GATTTTAATT TTTAAAAATG TATATTTTTT CCTGTGTAGA GTAGTAACAC CCTTGAAAAC 540

ACAWTCCCTT GTAAAGTCTC TAATTCTGTA CTCCGCATCT AGSTGRTCTC TTCTTTCTCA 600

GATATTTTAC AATTTCATTT ATCACCACCT TTCTCTAGCC TTTACCCGTC TCTTCAATAT 660

TWACATATGC AGAAGTTTCT CCTAACAAAC ACCTGCCTCT GCCTCAGTTC TGCTACCACC 720

CTGTTGCTTT CTTTCCCTTC ACAATCAAAT TTAAGAGTGT CAAAAAAAAA AAAAAAAAAC 780

TCGA 784

(2) INFORMATION FOR SEQ ID NO: 102:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1035 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 102:

AGAGGCCTGG CTGCGTTGCC CTATCTCCGT CTCCGCCACC CACTTAGCGT TTTAGGCATC 60

AATTACCAGC AGTTTCTCCG CCACTATCTG GAAAATTACC CGATTGCTCC CGGCAGAATA 120

CAAGAGCTTG AAGAACGCCG CAGTTGCGTG GAAGCCTGCA GAGCAAGGGA AGCAGCGTTT 180

GATGCCGAAT ATCAGCGAAA TCCTCACAGG GTGGACCTCG ATATTTTAAC CTTTACGATA 240

GCTCTGACTG CCTCTGAAGT TATCAACCCT CTGATAGAAG AACTTGGTTG CGATAAGTTT 300

ATCAATAGAG AATAGTTAGG TGGTGACACT ACTTCAAGAG AACCTCTGCA TTCCAGTCAT 360

ACCAATCCTG CAACTTGATT TTCAGAAGTC AAGAGTATAT CGCGATAAGA CAGTGCACAG 420

GTGGAGGGGA AAAAAAGGGG GAGGGGGAAG CTTATCTTGA AAAAGCATCA CAGAAGTAGA 480

AAAAAATGTC GAAAGCATTA TAACTGTAAC GTTCTTTGAG TTTGTGATTG ATCCACATTT 540

TTCCCCCTGC ATTATGGAAA ATGTCTCTCA GCATTGCTTT ATTACAAAGT AAAGGATGGT 600

TTTATAAAAT TGAGACTGAT GAAACATCAA TACTAGAGCC CATGAGGATG AAAGAAATTA 660

TCAAATAGTG CTGAACAGAA TAAGATGTTA ACGCTGAGTT ATTAGGACTG GAAGGCTATG 720

AAAAGAACTT GAAATTGTCG GAATATGTGC TCTCTTCATG TCATATTCAA TAGAAGTTTC 780

TAGTTTAAGA TTGATTTTGT GTTTTCTTAG GCATTTCAAG TGACAAGCAA AGTAAATGTA 840

TATATTATGT GATAAATCAT GTTTTCAAGA ACGTCAAATT TCTGGACTTT TTTCTTTCAA 900

TTTTTAATTT TTAAAGTTTT TTTGGTATTA AAAAATCYAT TCACAAGCCA AAAAATWTWT 960

WAAATWTWCM GCGAAAAGCC AAAAAAAAAA AAAAMMAGGG GGGGCCGGGC CCCATCCCCC 1020

CAAGGGGGTC CNGNT 1035

(2) INFORMATION FOR SEQ ID NO: 103:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2218 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 103:

AGGTATTAGG CCCTTTTGTG GGAGCCCCAT GTTTTGTTTT TCTGAGTTGG TGGGGAGGGA 60

SGGAGGGGGA GGGCTGAATT GTTTTGCAGA GGAAGATGGC ATCTGTGCTT TAAATTTCTC 120

ATTACTGGGT TAGAAAACAA AGAGGGAKTG CCCTGCACAT TTTCTTTTGT GCTTTTAAAT 180

GTTTCTTAAG TTGGAACAGG TTTCCTCGGG CCTGTTTTGA CTGATTGCTG GAGTGCATTT 240

GATAGTTAAA AATTACTAAT TGGTTTTATT TCCCTTCACA CTCTGCCTCC CCACTTCTCC 300

CCCCGTTACT GAAAAATAAC CATTTTAGTG TCAGGCTAGA AATTGAATTG CTGAGTTTTG 360

TGTATCCTTT AAATTAAAAA CCACAAGTGT TTATTGTAGT GGTTAAACTG TAGCATCTCA 420

GCATCTGGGT GGAAGCTGCC TATATTTCTT CCCAGTTTAA CTGGGGACCA TCTGTGAAAT 480

TAATTTTCCA TCCAGACAGC TGCTGTGAGC AAATGAACAT AAATGCTCGC TGGAAATTTA 540

CTAACCAGTT TTTATATTGA CCTGCAGTGT AAAAAGCACA TTTAATTATA AACAATATAT 600

TCAAAATGGG CAAATTTTAT TTTCAAATGC AGTGTAGAGC TAGATTAAAA GCAACTCTTT 660

GCCACCTACT CTGCCCTTTT GGCAAAGTTA CCTTGAACAA AGAATCTTAA GGGTTTATTA 720

AGAACTCTTT ATTTTCTTCA TACCCTGTTC TCTGCAGTGC TTTCTAACAG CTTCTGGGTG 780

CAGATTTTCT TCGGCATCCT TTTGCACTCA GCTTATTACA GGTAGGTAGT GCTTAAGAAA 840

AGTCATGGAG GACTAAAGCC TAAGTCCTTT TCACTTTTCC TCCATCTGAA GGTAGGTGAG 900

TTCATCCTCT TCATAGTAAT GCTGTTTTAC CAAGACTTTA TAGCAGATGG ACCCAGAAAG 960

AATTTTCTGC TATTGTGTTC ACTACAACAG GATAGGGACA TCAGACAGCC CCAGAAACCC 1020

CTTCCAGATC TGATATGGGA CTATTAATTT TTATGCTGTT AATTGGTATT CATTCACAAT 1080

GCAGTTGAAG GGGGAAGGCT CCACTGCATT CTTTGGCTAA GGCCTGAATG CTTGCTCATC 1140

TGTAAGATCT ATACTCGAGG TTTTGTTTTC CTTTTAAAAT TCTTTAGGGA GAGAGGGATG 1200

GTTTCTGAGG GGTTCTGAAA GTATGATTCA ATGTGCAACA TACAGGTAGG TCTTCAGCAT 1260

AAGCTGAAAT ATATGCATGT AAAAACTTTG ACATCTTTTT TTTTAATTTT CCACTTTCTT 1320

CTTAACTTTA CTTCTCTTTT TGTCCCCCCC CCATCTTACA GAAGTTGAGG CCAAGGGAGA 1380

ATGGTAGGCA CAGAAGAAAC ATGGCAAACT GCTCTGTGCT TTCAAACCAA AGTGTTCCCC 1440

CCAACCCCAA ATTTGTCTAA GCACTGGCCA GTCTGTTGTG GGCATTGTTT TCTACAACCA 1500

AATTCTGGGT TTTTTTCTTC TTTCTTTAAA CATAGAGGTA CCACCACAAG GGATGCCCTA 1560

CTCTCTCGCA GCTCTTGAAA GCATCTGTTT GAGGGAAAGG TCTCTGGGCA AGCAAGTGGT 1620

TATTTGGATT GCTTGCTTCC CTTTTTCCAC CTGGGACATT GYAATCATAA AATAACAGTA 1680

AATTCCAAAC CTCAAAAACT ATTATGGCCT GAGCACAGCT GAAATCTAGC AGAGTTTAAC 1740

TCTTCTGCCT CCATGTCTGT CACTTATAAT TCAGGTTCTG CTGTTGGCTT CAGAACATGA 1800

GCAGAAGAAT CGTTTTATGC TAGTTATTGC ATTCATGGTT GAAACTCAAC TTAGGGAAAG 1860

GGTTCCAATG TATTAAGCAA TGGGCTGCTT CTCCCCAATC CTCCCTAACA ATTCGTTGTG 1920

TGGACTTCTC ATCTAAAAGG TTAGTGGCTT TTGCTTGGGA TCAGTGCTCT CTATTGATGT 1980

TCTTGCTGGT CTCCAGACAC ATTCCTGTTG CATTAAGACT TGAAAGACTT GTAGATGTGT 2040

GATGTTCAGG CACAGGATGC TGAAAGCTAT GTTACTATTC TTAGTTTGTA AATTGTCCTT 2100

TTGATACCAT CATCTTGTTT TCTTTTTGTA GGTATAAATA AAAACACTGT TGACAATAAA 2160

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAA 2218

(2) INFORMATION FOR SEQ ID NO: 104:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1351 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 104:

CTTCACAGAC TGACAGAATG GTTTTGTTTT GTTTTGTTTT GTTTTGTTTT GTTTTTGAGA 60

TGGACTCTAG CTCTGTCACC CAGGCTGGAG TGCAGTGGTG CGATCTCGGC TCACTGCAAG 120

CTCCGCCTCC CGGGTTCTCA CCATTCTCCT GCCTCAGCCT CCCGAGTAGC TGGGACTACA 180

GGCGCCCACC ACCACGCCCG GCTAATTTTT TGTATTTTTT AGTAGAGACG GGGTTTCACC 240

ATGTTAGCCA GGATGGTCTC GATCTCCTGA CCTCGTGATC CGCCCGCYTC GGCCTCCCAA 300

AGTGCTGGGA TTACAGGCGT GAGCCACCGT GCCTGCCCCA GAATGGTTTT TAAAGCCACA 360

GTTGAGARGC CACCCATTGC CCGGCGCCTG GACAGTGATC ATCTTGTTCA TCTTGTTCAG 420

TCCTTTCTTG TGTGATTGGA ATTATTCATC CCCTTTGAAA GATGAGAAGG TTGAGATGCA 480

AAGAGTCTAC CTTTCCAAGT TCTCACTGCT GGAAAGARCT AGAAGCACAG TTCAAAGTTC 540

TGGNTTCTGG ACTCTGCAGT CCAGGTYTCC CTTYTCCCAC TTGCCTACCC TCAATGCCAC 600

ACTGTTTTTG AAGTGGCCCA TAACTTGAAG GRAAAGTTTA AAGACAGTTC AATTTAATCA 660

TCAGRATGCA TTCTTTTTTT TTTCGGARAC GGAKTTTCAC TCTTGCTGCC CASGCTGGAG 720

TGCAATGGTG CAATGATCTC GGCTCACTGC AACCTATGCC TCCTGGGTTC AAGNGATTAT 780

CCAGCCTCAG CCTCCCGAGT AGCTGGGATT ATGGGCGCCC ACCACCATGC CCAGCTAATT 840

TTTGTATTTT TTTTTTTAGT AGAGATGGGG TTTCGCCAGG TTGGCCAGGC TGKTCTTGTG 900

AAYTCCTGGC YTCAGGTGAT YTGCCCACYT CATCYTCCAA AAGTGCTGGG ATTACAGGCA 960

TGAGCCACTG CGCCTGGCYT CAGAATGCAT TCTTACACAT CTATCCTAGA CATTTATAAG 1020

CACTCTAATG GATAACAATC CAAGAATAAA TGATTGTAAA AGATGATGCC GAAGAGTTGA 1080

TGTCAATCTT TTTTTCCTAA GAAAAAAAGT CCGCGAGTAT TAAATATTTA GATCAATGTT 1140

TATAAAATGA TTACTTTGTA TATCTCATTA TTCCTATTTT GGAATAAAAA CTGACCTTCT 1200

TTAATCATAT ACTTGTCTTT TGTAAATAGC AGCTTTTGTG TCATTCTCCC CACTTTATTA 1260

GTTAATTTAA ATTGGAAAAA ACCCTCAAAC TAATATTCTT GTCTGTTCCA GTCTTATAAA 1320

TAAAACTTAT AATGCATGTA AAAAAAAAAA A 1351

(2) INFORMATION FOR SEQ ID NO: 105:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2066 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:

GGCACGAGGC GGCGGAGGGC CACAATCACA GCTCCGGGCA TTGGGGGAAC CCGAGCCGGC 60

TGCGCCGGGG GAATCCGTGC GGGCGCCTTC CGTCCCGGTC CCATCCTCGC CGCGCTCCAG 120

CACCTCTGAA GTTTTGCAGC GCCCAGAAAG GAGGCGAGGA AGGAGGGAGT GTGTGAGAGG 180

AGGGAGCAAA AAGCTCACCC TAAAACATTT ATTTCAAGGA GAAAAGAAAA AGGGGGGGCG 240

CAAAAATGGC TGGGGCAATT ATAGAAAACA TGAGCACCAA GAAGCTGTGC ATTGTTGGTG 300

GGATTCTGCT CGTGTTCCAA ATCATCGCCT TTCTGGTGGG AGGCTTGATT GCTCCAGGGC 360

CCACAACGGC AGTGTCCTAC ATGTCGGTGA AATGTGTGGA TGCCCGTAAG AACCATCACA 420

AGACAAAATG GTTCGTGCCT TGGGGACCCA ATCATTGTGA CAAGATCCGA GACATTGAAG 480

AGGCAATTCC AAGGGAAATT GAAGCCAATG ACATCGTGTT TTCTGTTCAC ATTCCCCTCC 540

CCCACATGGA GATGAGTCCT TGGTTCCAAT TCATGCTGTT TATCCTGCAG CTGGACATTG 600

CCTTCAAGCT AAACAACCAA ATCAGAGAAA ATGCAGAAGT CTCCATGGAC GTTTCCCTGG 660

CTTACCGTGA TGACGCATTT GCTGAGTGGA CTGAAATGGC CCATGAAAGA GTACCACGGA 720

AACTCAAATG CACCTTCACA TCTCCCAAGA CTCCAGAGCA TGAGGGCCGT TACTATGAAT 780

GTGATGTCCT TCCTTTCATG GAAATTGGGT CTGTGGCCCA TAAGTTTTAC CTTTTAAACA 840

TCCGGCTGCC TGTGAATGAG AAGAAGAAAA TCAATGTGGG AATTGGGGAG ATAAAGGATA 900

TCCGGTTGGT GGGGATCCAC CAAAATGGAG GCTTCACCAA GGTGTGGTTT GCCATGAAGA 960

CCTTCCTTAC GCCCAGCATC TTCATCATTA TGGTGTGGTA TTGGAGGAGG ATCACCATGA 1020

TGTCCCGACC CCCAGTGCTT CTGGAAAAAG TCATCTTTGC CCTTGGGATT TCCATGACCT 1080

TTATCAATAT CCCAGTGGAA TGGTTTTCCA TCGGGTTTGA CTGGACCTGG ATGCTGCTGT 1140

TTGGTGACAT CCGACAGGGC ATCTTCTATG CGATGCTTCT GTCCTTCTGG ATCATCTTCT 1200

GTGGCGAGCA CATGATGGAT CAGCACGAGC GGAACCACAT TGCAGGGTAT TGGAAGCAAG 1260

TCGGACCCAT TGCCGTTGGC TCCTTCTGCC TCTTCATATT TGACATGTGT GAGAGAGGGG 1320

TACAACTCAC GAATCCCTTC TACAGTATCT GGACTACAGA CATTGGAACA GAGCTGGCCA 1380

TGGCCTTCAT CATCGTGGCT GGAATCTGCC TCTGCCTCTA CTTCCTGTTT CTATGCTTCA 1440

TGGTATTTCA GGTGTTTCGG AACATCAGTG GGAAGCAGTC CAGCCTGCCA GCTATGAGCA 1500

AAGTCCGGCG GCTACACTAT GAGGGGCTAA TTTTTAGGTT CAAGTTCCTC ATGCTTATCA 1560

CCTTGGCCTG CGCTGCCATG ACTGTCATCT TCTTCATCGT TAGTCAGGTA ACGGAAGGCC 1620

ATTGGAAATG GGGCGGCGTC ACAGTCCAAG TGAACAGTGC CTTTTTCACA GGCATCTATG 1680

GGATGTGGAA TCTGTATGTC TTTGCTCTGA TGTTCTTGTA TGCACCATCC CATAAAAACT 1740

ATGGAGAAGA CCAGTCCAAT GGAATGCAAC TCCCATGTAA ATCGAGGGAA GATTGTGCTT 1800

TGTTTGTTTC GGAACTTTAT CAAGAATTGT TCAGCGCTTC GAAATATTCC TTCATCAATG 1860

ACAACGCAGC TTCTGGTATT TGAGTCAACA AGGCAACACA TGTTTATCAG CTTTGCATTT 1920

GCAGTTGTCA CAGTCACATT GATTGTACTT GTATACGCAC ACAAATACAC TCATTTAGCC 1980

TTTATCTCAA AATGTTAAAT ATAAGGAAAA AAGCGTCAAC AATAAATATT CTTGAGTATA 2040

AAAAAAAAAA AAAAAAAAAA AAAAAA 2066

(2) INFORMATION FOR SEQ ID NO: 106:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1705 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:

AATTCGGCAK AGGGCAGCTG TCGGCTGGAA GGAACTGGTC TGCTCACACT TGCTGGCTTG 60

CGCATCAGGA CTGGCTTTAT CTCCTGACTC ACGGTGCAAA GGTGCACTCT GCGAACGTTA 120

AGTCCGTCCC CAGCGCTTGG AATCCTACGG CCCCCACAGC CGGATCCCCT CAGCCTTCCA 180

GGTCCTCAAC TCCCGYGGAC GCTGAACAAT GGCCTCCATG GGGCTACAGG TAATGGGCAT 240

CGCGCTGGCC GTCCTGGGCT GGCTGGCCGT CATGCTGTGC TGCGCGCTGC CCATGTGGCG 300

CGTGACGGCC TTCATCGGCA GCAACATTGT CACCTCGCAG ACCATCTGGG AGGGCCTATG 360

GATGAACTGC GTGGTGCAGA GCACCGGCCA GATGCAGTGC AAGGTGTACG ACTCGCTGCT 420

GGCACTGCCG CAGGACCTGC AGGCGGCCCG CGCCCTCGTC ATCATCAGCA TCATCGTGGC 480

TGCTCTGGGC GTGCTGCTGT CCGTGGTGGG GGGCAAGTGT ACCAACTGCC TGGAGGATGA 540

AAGCGCCAAG GCCAAGACCA TGATCGTGGC GGGCGTGGTG TTCCTGTTGG CCGGCCTTAT 600

GGTGATAGTG CCGGTGTCCT GGACGGCCCA CAACATCATC CAAGACTTCT ACAATCCGCT 660

GGTGGCCTCC GGGCAGAAGC GGGAGATGGG TGCCTCGCTC TACGTCGGCT GGGCCGCCTC 720

CGGNCTGCTG CTCCTTGGCG GGGGGCTGCT TTGCTGCAAC TGTCCACCCC GCACAGACAA 780

GCCTTACTCC GCCAAGTATT CTGCTGCCCG CTCTGCTGCT GCCAGCAACT ACGTGTAAGG 840

TGCCACGGCT CCACTCTGTT CCTCTCTGCT TTGTTCTTCC CTGGACTGAG CTCAGCGCAG 900

GCTGTGACCC CAGGAGGGCC CTGCCACGGG CCACTGGCTG CTGGGGACTG GGGACTGGGC 960

AGAGACTGAG CCAGGCAGGA AGGCAGCAGC CTTCAGCCTC TCTGGCCCAC TCGGACAACT 1020

TCCCAAGGCC GCCTCCTGCT AGCAAGAACA GAGTCCACCC TCCTCTGGAT ATTGGGGAGG 1080

GACGGAAGTG ACAGGGTGTG GTGGTGGAGT GGGGAGCTGG CTTCTGCTGG CCAGGATGGC 1140

TTAACCCTGA CTTTGGGATC TGCCTGCATC GGTGTTGGCC ACTGTCCCCA TTTACATTTT 1200

CCCCACTCTG TCTGCCTGCA TCTCCTCTGT TGCGGGTAGG CCTTGATATC ACCTCTGGGA 1260

CTGTGCCTTG CTCACCGAAA CCCGCGCCCA GGAGTATGGC TGAGGCCTTG CCCACCCACC 1320

TGCCTGGGAA GTGCAGAGTG GATGGACGGG TTTAGAGGGG AGGGGCGAAG GTGCTGTAAA 1380

CAGGTTTGGG CAGTGGTGGG GGAGGGGGCC AGAGAGGCGG CTCAGGTTGC CCAGCTCTGT 1440

GGCCTCAGGA CTCTCTGCCT CACCCGCTTC AGCCCAGGGC CCCTGGAGAC TGATCCCCTC 1500

TGAGTCCTCT GCCCCTTCCA AGGACACTAA TGAGCCTGGG AGGGTGGCAG GGAGGAGGGG 1560

ACAGCTTCAC CCTTGGAAGT CCTGGGGTTT TTCCTCTTCC TTCTTTGTGG TTTCTGTTTT 1620

GTAATTTAAG AAGAGCTATT CATCACTGTA ATTATTATTA TTTTCTACAA TAAATGGGAC 1680

CTGTGCACAG GRAAAAAAAA AAAAG 1705

(2) INFORMATION FOR SEQ ID NO: 107:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1167 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:

TGCAGGAATT CGGCAGAGGT TTTCCGCTAG ACTCTGGCAG TTGGTGAGCA TCATGGCAAC 60

CGTTACAGCC ACAACCAAAG TCCCGGAGAT CCGTGATGTA ACAAGGATTG AGCGAATCGG 120

TGCCCACTCC CACATCCGGG GACTGGGGCT GGACGATGCC TTGGAGCCTC GGCAGGCTTC 180

GCAAGGCATG GTGGGTCAGC TGGCGGCACG GCGGGCGGCT GGCGTGGTGC TGGAGATGAT 240

CCGGGAAGGG AAGATTGCCG GTCGGGCAGT CCTTATTGCT GGCCAGCCGG GCACGGGGAA 300

GACGGCCATC GCCATGGGCA TGGCGCAGGC CCTGGGCCCT GACACGCCAT TCACAGCCAT 360

CGCCGGCAGT GAAATCTTCT CCCTGGAGAT GAGCAAGACC GAGGCGCTGA CGCAGGCCTT 420

CCGGCGGTCC ATCGGCGTTC GCATCAAGGA GGAGACGGAG ATCATCGAAG GGGAGGTGGT 480

GGAGATCCAG ATTGATCGAC CAGCAACAGG GACGGGCTCC AAGGTGGGCA AACTGACCCT 540

CAAGACCACA GAGATGGAGA CCATCTACGA CCTGGGCACC AAGATGATTG AKTCCCTGAC 600

CAAGGACAAG GTCCAGGCCG GGGACGTGAT CACCATCGAC AAGGCGACGG GCAAGATCTC 660

CAAGCTGGGC CGCTCCTTCA CACGCGCCCG CGAACTACGA CGCTATGGGC TCCCAGACCA 720

AGTTCGTGCA GTGCCCAGAT GGGGAGCTCC AGAAACGCAA GGAGGTGGTG CACACCGTGT 780

CCCTGCACGA GATCGACGTC ATCAACTCTC GCACCCAGGG CTTCCTGGCG CTCTTCTCAG 840

GTGACACAGG GGAGATCAAG TCAGAAGTCC GTGAGCAGAT CAATGCCAAG GTGGCTGAGT 900

GGCGCGAGGA GGGCAAGGCG GAGATCATCC CTGGAGTGCT GTTCATCGAC GAGGTCCACA 960

TGCTGGACAT CGAGAGCTTC TCCTTCCTCA ACCGGGCCCT GGAGAGTGAC ATGGCGCCTG 1020

TCCAGCAGGT CTATGGGGAT GCCGTGAGGG CTCTGGTAGC TGGTGCCCCG GATTCGCGTG 1080

ATGCCACGGT TGGTGGCCTC GTGCCGAATT CCTGCAGCCC GGGGGATCCA CTAGTTCTAG 1140

AGCGGCCGCC ACCGCGGTGG ANCTCCN 1167

(2) INFORMATION FOR SEQ ID NO: 108:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1907 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:

GGCACAGGGG AATCATCGTG TGATGTGTGT GCTGCCTTTG TGAGTGTGTG GAGTCCTGCT 60

CAGGTGTTAG GTACAGTGTG TTTGATCGTG GTGGCTTGAG GGGAACCCTT GTTCAGAGCT 120

GTGACTGCGG CTGCACTCAG AGAAGCTGCC CTTGGCTGCT CGTAGCGCCG GGCCTTCTCT 180

CCTCGTCATC ATCCAGAGCA GCCAGTGTCC GGGAGGCAGA AGGTACCGGG GCAGCTACTG 240

GAGGACTGTG CGGGCCTGCC TGGGCTGCCC CCTCCGCCGT GGGGCCCTGT TGCTGCTGTC 300

CATCTATTTC TACTACTCCC TCCCAAATGC GGTCGGCCCG CCCTTCACTT GGATGCTTGC 360

CCTCCTGGGC CTCTCGCAGG CACTGAACAT CCTCCTGGGC CTCAAGGGCC TGGCCCCAGC 420

TGAGATCTCT GCAGTGTGTG AAAAAGGGAA TTTCAACGTG GCCCATGGGC TGGCATGGTC 480

ATATTACATC GGATATCTGC GGCTGATCCT GCCAGAGCTC CAGGCCCGGA TTCGAACTTA 540

CAATCAGCAT TACAACAACC TGCTACGGGG TGCAGTGAGC CAGCGGCTGT ATATTCTCCT 600

CCCATTGGAC TGTGGGGTGC CTGATAACCT GAGTATGGCT GACCCCAACA TTCGCTTCCT 660

GGATAAACTG CCCCAGCAGA CCGGTGACCG TGCTGGCATC AAGGATCGGG TTTACAGCAA 720

CAGCATCTAT GAGCTTCTGG AGAACGGGCA GCGGGCGGGC ACCTGTGTCC TGGAGTACGC 780

CACCCCCTTG CAGACTTTGT TTGCCATGTC ACAATACAGT CAAGCTGGCT TTAGCGGGGA 840

GGATAGGCTT GAGCAGGCCA AACTCTTCTG CCGGACACTT GAGGACATCC TGGCAGATGC 900

CCCTGAGTCT CAGAACAACT GCCGCCTCAT TGCCTACCAG GAACCTGCAG ATGACAGCAG 960

CTTCTCGCTG TCCCAGGAGG TTCTCCGGCA CCTGCGGCAG GAGGAAAAGG AAGAGGTTAC 1020

TGTGGGCAGC TTGAAGACCT CAGCGGTGCC CAGTACCTCC ACGATGTCCC AAGAGCCTGA 1080

GCTCCTCATC AGTGGAATGG AAAAGCCCCT CCCTCTCCGC ACGGATTTCT CTTGAGACCC 1140

AGGGTCACCA GGCCAGAGCC TCCAGTGGTC TCCAAGCCTC TGGACTGGGG GCTCTCTTCA 1200

GTGGCTGAAT GTCCAGCAGA GCTATTTCCT TCCACAGGGG GCCTTGCAGG GAAGGGTCCA 1260

GGACTTGACA TCTTAAGATG CGTCTTGTCC CCTTGGGCCA GTCATTTCCC CTCTCTGAGC 1320

CTCGGTGTCT TCAACCTGTG AAATGGGATC ATAATCACTG CCTTACCTCC CTCACGGTTG 1380

TTGTGAGGAC TGAGTGTGTG GAAGTTTTTC ATAAACTTTG GATGCTAGTG TACTTAGGGG 1440

GTGTGCCAGG TGTCTTTCAT GGGGCCTTCC AGACCCACTC CCCACCCTTC TCCCCTTCCT 1500

TTGCCCGGGG ACGCCGAACT CTCTCAATGG TATCAACAGG CTCCTTCGCC CTCTGGCTCC 1560

TGGTCATGTT CCATTATTGG GGAGCCCCAG CAGAAGAATG GAGAGGAGGA GGAGGCTGAG 1620

TTTGGGGTAT TGAATCCCCC GGCTCCCACC CTGCAGCATC AAGGTTGCTA TGGACTCTCC 1680

TGCCGGGCAA CTCTTGCGTA ATCATGACTA TCTCTAGGAT TCTGGCACCA CTTCCTTCCC 1740

TGGCCCCTTA AGCCTAGCTG TGTATCGGCA CCCCCACCCC ACTAGAGTAC TCCCTCTCAC 1800

TTGCGGTTTC CTTATACTCC ACCCCTTTCT CAACGGTCCT TTTTTAAAGC ACATCTCAGA 1860

TTAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAGGG CGGCCGC 1907

(2) INFORMATION FOR SEQ ID NO: 109:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 611 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 109:

ATGAATTAAC GCCAAGCTNT NAATAGGGAC TCACTATGGG GGAAAGNTGG GTAACGCCTG 60

CAGGTACCGT TCCGGAATTC CCGGGTCGAC CCACGCGTCC GATGGGGCTT TAGTAAATCA 120

GGCTTGCAGG CTCAAAGCTG CAATCTGCCC ACTCTCAGGT ACTGAGACTT TGTGGGCCTC 180

AGACACCAGG AAGAAAGTTG GGATACAGTC ATTTGAGTTA AAAAGGGAAT GACCCCTCAG 240

AAACCCGCAT TAGCAGTGTT ACTCTTGGAA GTGCCTTTAC TTTTAACGCT CTCTGTTCTG 300

AAAAAGAGGT GTTTGGTTAC GTGTGAGCCA ACATCACGTT TTGTTAGCTG TGATTTACCT 360

TTGTCCGTTT AAAAGACTTC ACGGAGCCAT TCTGTATACA AGGTGTGCTC TTTCCAATGT 420

AGAAGGGGTT ATGGAAAAGG GTGCGATCCT TTGCTGTAAA CTGGAGAGAC CAGTCCCAAA 480

CAGAGGGGAA TTTTAAGCCC TTCTCATCAC CCAATTGGAT GTTTTTGCTT ATAGCAAATT 540

CCTGCAAAAT AAATAAATAA ATATTTGCAA AACTAAAAAA AAAAAAAAAA AAAAAAAAAA 600

GGGGGGNCCN C 611

(2) INFORMATION FOR SEQ ID NO: 110:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2632 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 110:

TCCCAGCTCT CAGGACAAGG GCCCTGGGCG ATCTTTTAAA AAAGCCGATT GGGTGTCTTT 60

CTAAAANTAC AACCAGTACT TCATCGTCAA GTTTCTGGGA AGGGAGTCCC CTCCAGATTC 120

TCATGGAGTG ACAAATCTTG ACTCTTGCTC CTGGAATTTT TCAGGCCCAA ACTAGCGTTT 180

CTACAATGAT TTATTTGGCA AATTTGTCTT GATTATGGGT GGCTGATGAG GAACGTGCTT 240

TTGTTAGGAA CCGAAACTGG GCGGCGGTGA GGGCGTGTAC GCAATGAGTC CGGAAGAGGG 300

TGAAATGCTT TCGGTAGGCA CTCCACGGCT GTGAAGATGG CGGCGGCTGC GTGGCTTCAG 360

GTGTTGCCTG TCATTCTTCT GCTTCTGGGA GCTCACCCGT CACCACTGTC GTTTTTCAGT 420

GCGGGACCGG CAACCGTAGC TGCTGCCGAC CGGTCCAAAT GGCACATTCC GATACCGTCG 480

GGGAAAAATT ATTTTAGTTT TGGAAAGATC CTCTTCAGAA ATACCACTAT CTTCCTGAAG 540

TTTGATGGAG AACCTTGTGA CCTGTCTTTG AATATAACCT GGTATCTGAA AAGCGCTGAT 600

TGTTACAATG AAATCTATAA CTTCAAGGCA GAAGAAGTAG AGTTGTATTT GGAAAAACTT 660

AAGGAAAAAA GAGGCTTGTC TGGGAAATAT CAAACATCAT CAAAATTGTT CCAGAACTGC 720

AGTGAACTCT TTAAAACACA GACCTTTTCT GGAGATTTTA TGCATCGACT GCCTCTTTTA 780

GGAGAAAAAC AGGAGGCTAA GGAGAATGGA ACAAACCTTA CCTTTATTGG AGACAAAACC 840

GCAATGCATG AACCATTGCA AACTTGGCAA GATGCACCAT ACATTTTTAT TGTACATATT 900

GGCATTTCAT CCTCAAAGGA ATCATCAAAA GAAAATTCAC TGAGTAATCT TTTTACCATG 960

ACTGTTGAAG TGAAGGGTCC CTATGAATAC CTCACACTTG AAGACTATCC CTTGATGATT 1020

TTTTTCATGG TGATGTGTAT TGTATATGTC CTGTTTGGTG TTCTGTGGCT GGCATGGTCT 1080

GCCTGCTACT GGAGAGATCT CCTGAGAATT CAGTTTTGGA TTGGTGCTGT CATCTTCCTG 1140

GGAATGCTTG AGAAAGCTGT CTTCTATGCG GAATTTCAGA ATATCCGATA CAAAGGARAA 1200

TCTGTCCAGG GTGCTTTGAT CCTTGCAGAR CTGCTTTCAG CAGTGAAACG CTCACTGGCT 1260

CGAACCCTGG TCATCATAGT CAGTCTGGGA TATGGCATCG TCAAGCCACG CCTGGAGTCA 1320

CTCTTCATAA GGTTGTAGTA GCAGRAGCCC TCTATCTTTT GTTCTCTGGC ATGGAAGGGG 1380

TCCTCAGAGT TACTGGGGCC CAGACTGATC TTGCTTCCTT GGCCTTTATC CCCTTGGCTT 1440

TCCTAGACAC TGCCTTGTGC TGGTGGATAT TTATTAGCCT GACTCAAACA ATGAAGCTAT 1500

TAAAACTTCG GAGGAACATT GTAAAACTCT CTTTGTATCG GCATTTCACC AACACGCTTA 1560

TTTTGGCAGT GGCAGCATCC ATTGTGTTTA TCATCTGGAC AACCATGAAG TTCAGAATAG 1620

TGACATGTCA GTCGGACTGG CGGGAGCTGT GGGTAGACGA TGCCATCTGG CGCTTGCTGT 1680

TCTCCATGAT CCTCTTTGTC ATCATGGTTC TCTGGCGACC ATCTGCAAAC AACCAGAGGT 1740

TTGCCTTTTC ACCATTGTCT GAGGAAGAGG AGGAGGATGA ACAAAAGGAG CCTATGCTGA 1800

AAGAAAGCTT TGAAGGAATG AAAATGAGAA GTACCAAACA AGAACCCAAT GGAAATAGTA 1860

AAGTTAACAA AGCACAGGAA GATGATTTGA AGTGGGTAGA AGAGAATGTT CCTTCTTCTG 1920

TGACAGATGT AGCACTTCCA GCCCTTCTGG ATTCAGATGA GGAACGAATG ATCACACACT 1980

TTGAAAGGTC CAAAATGGAG TAAGGAATGG GAAGATTTGC AGTTAAAGAT GGCTACCATC 2040

AGGGAAGAGA TCAGCATCTG TGTCAGTCTT CTGTACGGCT CCATGGGATT AAAGGAAGCA 2100

ATGACATCCT GATCTGTTCC TTGATCTTTG GGCATTGGAG TTGGCGAGAG GTGTCAGAAC 2160

AAAGAGAACA TCTTACTGAA AACAAGTTCA TAAGATGAGA AAAATCTACG AGCTTCTTAT 2220

TTACAACACT GCTGCCCCCT TTCCTCCCAG ACTCTGACAT GGATGTTCAT GCAACTTAAG 2280

TGTGTTGTTC CTGAACTTTC TGTAATGTTT CATTTTTTAA ATCTGACAAA CTAAAAAGTT 2340

TAACGTCTTC TAAAAGATTG TCATCAACAC CATAATATGT AATCTCCAGG AGCAACTGCC 2400

TGTAATTTTT ATTTATTTAG GGAGTTACAT AGGTGATGGG GGAAATTGTT AACTACCTTT 2460

CATTTTCCTG GGAAGTCAAG GTTACATCTT GCAGAGGTTG TTTTGAGAAA AAAGGGCCCT 2520

TCTGAGTTAA GGAGCCATAG TTCTATCAAT GATCAAAAGA AAAAAAAAAA AACTCGATCG 2580

GCACGAGGGG GGGCCCGGTA CCCAATTCGC CCTATGGGAN TCGAATGAGA CC 2632

(2) INFORMATION FOR SEQ ID NO: 111:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2249 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:

GAATTCGGCA CGAGCTCACC GTGCTGCGTG ACACAAGGCC AGCCTGCGCC TACGAGCCCA 60

TGGACTTTKT RATGGCCCTC ATCTACGACA TGGTACTGSW TGTGGTCACC CTGGGGCTGG 120

CCCTCTTCAC TCTGTGCGGC AAGTTCAAGA GGTGGAAGCT GAACGGGGCC TTCCTCCTCA 180

TCACAGCCTT CCTCTCTGTG CTCATCTGGG TGGCCTGGAT GACCATGTAC CTCTTCGGCA 240

ATGTCAAGCT GCAGCAGGGG GATGCCTGGA ACGACCCCAC CTTGGCCATC ACGCTGGCGG 300

CCAGCGCTGG GTCTTCGTCA TCTTCCACGC CATCCCTGAG ATCCACTGCA CCCTTCTGCC 360

AGCCCTGCAG GAGAACACGC CCAACTACTT CGACACGTCG CAGCCCAGGA TGCGGGAGAC 420

GGCCTTCGAG GAGGACGTGC AGCTGCCGCG GGCCTATATG GAGAACAAGG CCTTCTCCAT 480

GGATGAACAC AATGCAGCTC TCCGAACAGC AGGATTTCCC AACGGCAGCT TGGGAAAAAG 540

ACCCAGTGGC AGCTTGGGGA AAAGACCCAG CGCTCCGTTT AGAAGCAACG TGTATCAGCC 600

AACTGAGATG GCCGTCGTGC TCAACGGTGG GACCATCCCA ACTGCTCCGC CAAGTCACAC 660

AGGAAGAMAC CTTTGGTGAA AGACTTTAAG TTCCAGAGAA TCAGAATTTC TCTTACCGAT 720

TTGCCTCCCT GGCTGTGTCT TTCTTGAGGG AGAAATCGGT AACAGTTGCC GAACCAGGCC 780

GCCTCACAGC CAGGAAATTT GGAAATCCTA GCCAAGGGGA TTTCGTGTAA ATGTGAACAC 840

TGACGAACTG AAAAGCTAAC ACCGACTGCC CGCCCCTCCC CTGCCACACA CACAGACACG 900

TAATACCAGA CCAACCTCAA TCCCCGCAAA CTAAAGCAAA GCTAATTGCA AATAGTATTA 960

GGCTCACTGG AAAATGTGGC TGGGAAGACT GTTTCATCCT CTGGGGGTAG AACAGAACCA 1020

AATTCACAGC TGGTGGGCCA GACTGGTGTT GGTTGGAGGT GGGGGGCTCC CACTCTTATC 1080

ACCTCTCCCC AGCAAGTGCT GGACCCCAGG TAGCCTCTTG GAGATGACCG TTGCGTTGAG 1140

GACAAATGGG GACTTTGCCA CCGGCTTTGC CTGGTGGTTT GCACATTTCA GGGGGGTCAG 1200

GAGAGTTAAG GAGGTTGTGG GTGGGATTCC AAGGTGAGGC CCAACTGAAT CGTGGGGTGA 1260

GCTTTATAGC CAGTAGAGGT GGAGGGACCC TGGCATGTGC CAAAGAAGAG GCCCTCTGGG 1320

TGATGAAGTG ACCATCACAT TTGGAAAGTG ATCAACCACT GTTCCTTCTA TGGGGCTCTT 1380

GCTCTAGTGT CTATGGTGAG AACACAGGCC CCGCCCCTTC CCTTGTAGAG CCATAGAAAT 1440

ATTCTGGCTT GGGGCAGCAG TCCCTTCTTC CCTTGATCAT CTCGCCCTGT TCCTACACTT 1500

ACGGGTGTAT CTCCAAATCC TCTCCCAATT TTATTCCCTT ATTCATTTCA AGAGCTCCAA 1560

TGGGGTCTCC AGCTGAAANS CCCTCCGGGA GGCAGGTTGG AAGGCAGGCA CCACGGCAGG 1620

TTTTCCGCGA TGATGTCACC TAGCAGGGCT TCAGGGGTTC CCACTAGGAT GCAGAGATGA 1680

CCTCTCGCTG CCTCACAAGC AGTGACACCT CGGGTCCTTT CCGTTGCTAT GGTGAAAATT 1740

CCTGGATGGA ATGGATCACA TGAGGGTTTC TTGTTGCTTT TGGAGGGTGT GGGGGATATT 1800

TTGTTTTGGT TTTTCTGCAG GTTCCATGAA AACAGCCCTT TTCCAAGCCC ATTGTTTCTG 1860

TCATGGTTTC CATCTGTCCT GAGCAAGTCA TTCCTTTGTT ATTTAGCATT TCGAACATCT 1920

CGGCCATTCA AAGCCCCCAT GTTCTCTGCA CTGTTTGGCC AGCATAACCT CTAGCATCGA 1980

TTCAAAGCAG AGTTTTAACC TGACGGCATG GAATGTATAA ATGAGGGTGG GTCCTTCTGC 2040

AGATACTCTA ATCACTACAT TGCTTTTTCT ATAAAACTAC CCATAAGCCT TTAACCTTTA 2100

AAGAAAAATG AAAAAGGTTA GTGTTTGGGG GCCGGGGGAG GACTGACCGC TTCATAAGCC 2160

AGTACGTCTG AGCTGAGTAT GTTTCAATAA ACCTTTTGAT ATTTCTCAAA AAAAAAAAAA 2220

AAAAANCCCG GGGGGGGGCC CGGACCTGG 2249

(2) INFORMATION FOR SEQ ID NO: 112:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2198 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:

GATACTATAA GGCAAGTGAC TCACGGGTGC GCCGTTAGAC TAGTGGATCC CGGGTGCAGG 60

AATTCGGCAG AGCGCCGCCG GAGCCGAAGT GCTGGCGCCC CCGCGGCCGC TGCCTCCGCG 120

GANCCCAAAA TCATGAAAGT CACCGTGAAG ACCCCGAAGA AAAGGAGGAA TTCGCCGTGC 180

CCGAGAATAG CTCCGTCCAG CAGTTTAAGG AAGAAATCTC TAAACGTTTT AAATCACATA 240

CTGACCAACT TGTGTTGATA TTTGCTGGAA AAATTTTGAA AGATCAAGAT ACCTTGAGTC 300

AGCATGGAAT TCATGATGGA CTTACTGTTC ACCTTGTCAT TAAAACACAA AACAGGCCTC 360

AGGATCATTC AGCTCAGCAA ACAAATACAG CTGGAAGCAA TGTTACTACA TCATCAACTC 420

CTAATAGTAA CTCTACATCT GGTTCTGCTA CTAGCAACCC TTTTGGTTTA GGTGGCCTTG 480

GGGGACTTGC AGGTCTGAGT AGCTTGGGTC TGAATACTAC CAACTTCTCT GAACTACAGA 540

GTCAGATGCA GCGACAACTT TTGTCTAACC CTGAAATGAT GGTCCAGATC ATGGAAAAWC 600

CCYTTGTTCA GAGCATGCTC NTCAAATCCT GACCTGATGN AGACAGTTAA TTATGGCCAA 660

TCCACAAATG CAGCAGTTGA TACAGAGAAA TCCCAGAAAT TAGTCATATG TTGAATAATC 720

CAGATATAAT GAGACAAACG TTGGAACTTG CCCAGGAATC CAGCAATGAT GCAGGAGATG 780

ATGAGGAACC AGGACCGAGC TTTGAGCAAC CTAGAAAGCA TCCCAGGGGG ATATAATGCT 840

TTAAGGCGCA TGTACACAGA TATTCAGGAA CCAATGCTGA GTGCTGCACA AGAGCAGTTT 900

GGTGGTAATC CATTTGCTTC CTTGGTGAGC AATACATCCT CTGGTGAAGG TAGTCAACCT 960

TCCCGTACAG AAAATAGAGA TCCACTACCC AATCCATGGG CTCCACAGAC TTCCCAGAGT 1020

TCATCAGCTT CCAGCGGCAC TGCCAGCACT GTGGGTGGCA CTACTGGTAG TACTGCCAGT 1080

GGCACTTCTG GGCAGAGTAC TACTGCGCCA AATTTGGTGC CTGGAGTAGG AGCTAGTATG 1140

TTCAACACAC CAGGAATGCA GAGCTTGTTG CAACAAATAA CTGAAAACCC ACAACTTATG 1200

CAAAACATGT TGTCTGCCCC CTACATGAGA AGCATGATGC AGTCACTAAG CCAGAATCCT 1260

GACCTTGCTG CACAGATGAT GCTGAATAAT CCCCTATTTG CTGGAAATCC TCAGCTTCAA 1320

GAACAAATGA GACAACAGCT CCCAACTTTC CTCCAACAAA TGCAGAATCC TGATACACTA 1380

TCAGCAATGT CAAACCCTAG AGCAATGCAG GCCTTGTTAC AGATTCAGCA GGGTTTACAG 1440

ACATTAGCAA CGGAAGCCCC GGGCCTCATC CCAGGGTTTA CTCCTGGCTT GGGGGCATTA 1500

GGAAGCACTG GAGGCTCTTC GGGAACTAAT GGATCTAACG CCACACCTAG TGAAAACACA 1560

AGTCCCACAG CAGGAACCAC TGAACCTGGA CATCAGCAGT TTATTCAGCA GATGCTGCAG 1620

GCTCTTGCTG GAGTAAATCC TCAGCTACAG AATCCAGAAG TCAGATTTCA GCAACAACTG 1680

GAACAACTCA GTGCAATGGG ATTTTTGAAC CGTGAAGCAA ACTTGCAAGC TCTAATAGCA 1740

ACAGGAGGTG ATATCAATGC AGCTATTGAA AGGTTACTGG GCTCCCAGCC ATCATAGCAG 1800

CATTTCTGTA TCTKGAAAAA ATGTAATTTA TTTTTGATAA CGGCTCTTAA ACTTTAAAAT 1860

ACCTGCTTTA TTTCATTTTG ACTCTTGGAA TTCTGTGCTG TTATAAACAA ACCCAATATG 1920

ATGCATTTTA AGGTGGAGTA CAGTAAGATG TGTGGGTTTT TCTGTATTTT TCTTTTCTGG 1980

AACAGTGGGA ATTAAGGCTA CTGCATGCAT CACTTCTGCA TTTATTGTAA TTTTTTAAAA 2040

ACATCACCTT TTATAGTTGG GTGACCAGAT TTTGTCCTGC ATCTGTCCAG TTTATTTGCT 2100

TTTTAAACAT TAGCCTATGG TAGTAATTTA TGTAGAATAA AAGCATTAAA AAGAAGCAAA 2160

AAAAAAAAAA AAAAATTCCT GCGCCCGCGA ATTCTTCT 2198

(2) INFORMATION FOR SEQ ID NO: 113:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1043 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 113:

CTGAAGTGTA TGTGGTGAGG AAGAAGAGGC TCCTACTGTA GACAGCCTTG TTCTACAGAT 60

CCTCCCAGAA ATCTCTGGGC CAGGTGGAAC CCAGGGTCAG AGAGGGATGG GAGAGAGGTT 120

TAATTTTCCA TGATAAATAA AAATCTATAA AATAATAAAC AAGAGAAAAG AGATTGGAAA 180

CAGCCAGGTT GGAGCAGTGA GTGAGTAAGG AAACCTGGCT GCCCTCTCCA GATTCCCCAG 240

GCTCTCAGAG AAGATCAGCA GAAAGTCTGC AAGACCCTAA GAACCATCAG CCCTCAGCTG 300

CACCTCCTCC CCTCCAAGGA TGACAAAGGC GCTACTCATC TATTTGGTCA GCAGCTTTCT 360

TGCCCTAAAT CAGGCCAGCC TCATCAGTCG CTGTGACTTG GCCCAGGTGC TGCAGCTGGA 420

RGACTTGGAT GGGTTTGAGG GTTACTCCCT GAGTGACTGG CTGTGCCTGG CTTTTGTGGA 480

AAGCAAGTTC AACATATCAA AGATWAATGA AAATGCAGAT GGAAGCTTTG ACTATGGSCT 540

CTTCCAGATC AACAGCCACT ACTGGTGCAA CRATTATAAG AGTTACTCGG AAAACCTTTG 600

CCACGTAGAC TGTCAAGATC TGCTGAATCC CAACCTTCTT GCAGGCATCC ACTGCGCAAA 660

AAGGATTGTG TCCGGAGCAC GGGGGATGAA CAACTGGGTT AGAATGGAAG KTTGCACTGT 720

TCAGGCCGGC CACTCTTCTA CTGGCTGACA GGATGCCGCC TGAGATKAAA CARGGTGCGG 780

GTGCACCGTG GARTCATTCC AAGACTCCTG TCCTCACTCA RGGATTCTTC ATTTCTTCTT 840

CCTACTGCCT CCACTTCATG TTATTTTCTT CCCTTCCCAT TTACAACTAA AACTGACCAG 900

AGCCCCAGGA ATAAATGGTT TTCTTGGCTT CCTCCTTACT CCCATCTGGA CCCAGTCCCC 960

TGGTTCCTGT CTGTTATTTG TAAACTGAGG ACCACAATAA AGAAATCTTT ATATTTATCG 1020

AAAAAAAAAA AAAAAAAACT CGA 1043

(2) INFORMATION FOR SEQ ID NO: 114:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 703 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 114:

GAATTCGGCA CGAGTGCGCG GGCACCACGG CGGTTTTTCG ACGCTGGCGG TGGACGCAGG 60

CAGCATGGAC CACGGTTGCT GGGCGGATGG GGAGCGTCTA TGGTCAGTTG CCTTAGAAGT 120

GGTGAGATGG GAAGCTGCAG TTGGAAGACC CTGGAGGATG CCTGACAAGG GGATGTCTGA 180

CACATGATTG GAGCTCTTTT TGAAATGTTT CTTGCCCTTC CTGGAGCAGA GGAGCCATTA 240

TTTATGCAGG TACATCGAAG TCTTTTGACC TCCATACAGT GATTATGCTT GTCATCGCTG 300

GTGGTATCCT GGCGGCCTTG CTCCTGCTGA TAGTTGTCGT GCTCTGTCTT TACTTCAAAA 360

TACACAACGC GCTAAAAGCT GCAAAGGAAC CTGAAGCTGT GGCTGTAAAA AATCACAACC 420

CAGACAAGGT GTGGTGGGCC AAGAACAGCC AGGCCAAAAC CATTGCCACG GAGTCTTGTC 480

CTGCCCTGCA GTGCTGTGAA GGATATAGAA TGTGTGCCAG TTTTGATTCC CTGCCACCTT 540

GCTGTTGCGA CATAAATGAG GGCCTCTGAG TTAGGAAAGG TGGGCACAAA AATCTTCATG 600

AGCAATACTT CTTAGTAGAT TGTTTTGTTA TTCAAATCAA GTTCTAGTGT TTTTATGTGA 660

GATTATATAA TTTACAGTGT TGTTTTATAT ACTTTTGAAT AAA 703

(2) INFORMATION FOR SEQ ID NO: 115:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3684 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 115:

GGCAGAGGGG GCATGAGCAG GAGGAGGATT ACCGCTACGA GGTGCTCACG GCCGAGCAGA 60

TTCTACAACA CATGGTGGNA ATGTATCCGG GAGGTCAACG AGGTCATCCA GAATCCAGCA 120

ACTATCACAA GAATACTCCT TAGCCACTTC AATTGGGATA AAGAGAAGCT AATGGAAAGG 180

TACTTTGATG GAAACCTGGA GAAGCTCTTT GCTGAGTGTC ATGTAATTAA TCCAAGTAAA 240

AAGTCTCGAA CACGCCAGAT GAATACAAGG TCATCAGCAC AGGATATGCC TTGTCAGATC 300

TGCTACTTGA ACTACCCTAA CTCGTATTTC ACTGGCCTTG AATGTGGACA TAAGTTTTGT 360

ATGCAGTGCT GGAGTGAATA TTTAACTACC AAAATAATGG AAGAAGGCAT GGGTCAGACT 420

ATTTCGTGTC CTGCTCATGG TTGTGATATC TTAGTGGATG ACAACACAGT TATGCGCCTG 480

ATCACAGATT CAAAAGTTAA ATTAAAGTAT CAGCATTTAA TAACAAATAG CTTTGTAGAG 540

TGCAATCGAC TGTTAAAGTG GTGTCCTGCC CCAGATTGCC ACCATGTTGT TAAAGTCCAA 600

TATCCTGATG CTAAACCTGT TCGCTGCAAA TGTGGGCGCC AATTTTGCTT TAACTGTGGA 660

GAAAATTGGC ATGATCCTGT TAAATGTAAG TGGTTAAAGA AATGGATTAA AAAGTGTGAT 720

GATGACAGTG AAACCTCCAA TTGGATTGCA GCCAACACAA AGGAATGTCC CAAATGCCAT 780

GTCACAATTG AGAAGGATGG TGGTTGTAAT CACATGGTCT GTCGTAACCA GAATTGTAAA 840

GCAGAGTTTT GCTGGGTGTG TCTTGGCCCA TGGGAACCAC ATGGATCTGC CTGGTACAAC 900

TGTAACCGCT ATAATGAGGA TGATGCAAAG GCAGCAAGAG ATGCACAGGA GCGATCTAGG 960

GCAGCCCTGC AGAGGTACCT GTTCTACTGT AATCGCTATA TGAACCACAT GCAGAGCCTG 1020

CGCTTTGAGC ACAAACTATA TGCTCAGGTG AAACAGAAAA TGGAGGAGAT GCAGCAGCAC 1080

AACATGTCCT GGATTGAGGT GCAGTTCCTG AAGAAGGCAG TTGATGTCCT CTGCCAGTGT 1140

CGTGCCACAC TCATGTACAC TTATGTCTTC GCTTTCTACC TCAAAAAGAA TAACCAGTCC 1200

ATTATCTTTG AGAATAACCA AGCAGATCTA GAGAATGCCA CAGAGGTGCT CTCGGGCTAC 1260

CTTGAACGAG ATATTTCCCA AGATTCTCTG CAGGATATAA AGCAGAAAGT ACAAGACAAG 1320

TACAGATACT GTGAGAGTCG ACGAAGGGTT TTGTTACAGC ATGTGCATGA AGGCTATGAA 1380

AAAGATCTGT GGGAGTACAT TGAGGACTGA GAATGGCCCT GCATAAAATG AACTCTGAAA 1440

ACTTTACCAT CTAGAGTGCT CATGCAATTA AAACAAAACA AACACAAACA AGGAGGCACT 1500

AAGCCTATTC TGACACCACT GGTCTGTAGT ACCAGAATTG TTTTGTTAAT GGAAAGTTTA 1560

AGTAAATTAT ATTGTAATAA AAAGGTAGAT AAACCATTGT ACAACAGTAT TCTAGGCCGC 1620

CAACAAAAGT GTGACAGACA CACTAAAAGC CCTCCAACTT TAACTTGTAA CGTAGCTTCA 1680

TTCTCAAAGC TGACTCCTTT TTTTTCTTTT TCCTTTTCCT GAGTGTAGTA CAGTTAAAAT 1740

TTCAAACAGC TCCTTGACAC TGCTTTTCAT GTTCAAACCA GCCATTTTGT TGTACTTTGG 1800

TAAAGGACCT CTTCCCCTTC CTCCCCTACA CATACAGATA CACCCACACA CAGACTGACT 1860

CTCTTTCTCT CATACCCCAA GGTCATGAGT GAATGATGCT TAGTTCCTTG TAAAGAAAAT 1920

CTTGGGATGG GGAAAGGGGT AGGCAGCAAG AGGATTCAAC AAACGAAAAA CATAAAAACT 1980

TTGTATATGA CTTTTAAAAC AAGAGGACAA CACAGTATTT TTCAAAATTG TATATAGCGC 2040

ATATGCATGG ACAAAGCAAG CGTGGCACGT GTTTGCATAA TGTTTAATTA CAAAAAAATA 2100

TTTATTCTTT AAAAATCTTC AAGATTATGT CTATTTGCTG TGCATTTTCT TTCAGTTTGC 2160

TTATCTTTCC CGGGTTGGGG TTGGGATAAA GGTGTGTCGG TTTAGCACCT CTGGAAGACC 2220

TATCTAGAGC TCTTTCACTT TCCTGAGGTT ATTTTGCCCY TTCTGGTGTT GGTATGTCTG 2280

TTGCCGGCCA TGGGCTNCAY GCCTTGAATT CCTGCTCTTG ATCAGGGACA AGGGAGGTCA 2340

AGCTCTGACT AATGCCATGA CCTGATTAAG GGGTACAGCA GGGAGTTTTG TTGCTACAGC 2400

TCATGAATTA ACCTGTCCCA ACCTAATCCC CCTCCATGGC ATCATGCCTC TACCCAAGCC 2460

TTTGTGTGCC CATGTTATGC ACACAGCTGT AGGCATTCTT AAGTCCCCTG TCGCATCCAG 2520

TGGAAGCATT TTAAAATTTC TTTTACTTTT TGGTTTTCCC TTAATTGCTG CTTTTCAGAT 2580

TTTAGTTATG GCTCGTCTGC TCACCCCTTC TCTACATTAG GGTGTCAAAG AGAATGTTTT 2640

GCTTTAAATA TAAATAGCCA TTCATTTAGT CTCAGATTGT GAATTTAAAA TGGTGGATAC 2700

CGAAATTGCT TGTGTGTGTT GCTGTGGGTT TGGTTTGAAG GCAAACACCC CTAGAACATG 2760

ATATTCCCAT CTAGTGCATT TAAATAGAAA TCACTGAGTT TGCTGCTTTT TTATTGTCAG 2820

CAGATAGGAG AATTAATAAT GCATTTTAGC TGTGATGTCC ATTTTTATGA AATTCCTACT 2880

AAGAGCTATG TTAAAAGTAA AGGATGGTGG TGGTTGTATT AACTATATAC CTGTTTAGGC 2940

CATTCTGGCT GTGGTATTTT TCAATAGGTC AGCATCTGTA AATCTGTCAG TTTTATACAG 3000

GAGTGCAGAG TGAACTAGGC AACTAGATTA AGAGGTCTAA ATATGAAATA CCAGTTGAGG 3060

CTGAGGACCT CTTCGTCTTC CTTTAAATGT CTTTTGCCTA GGGAGTGTTT ACCATTTGTG 3120

AGGCAGCTTT GTCTGCTCTT ACACTGTACA TCCTATTACT CCATTGGGAA GTAGGTTCAC 3180

TTTCCTCTGG CCTTTTGCCT AAGTTAGGCT TTGCTGAATC AACCCTACTT TTCCTTTTAG 3240

AAAAGGTTGT TACAGGAGAT TTACTGGCAA CTGTTCTTTT CCCATCAAAA ATCAGTGAAT 3300

GTTTGCTGAG TATAAATGCT GCTTCCTTAA ACCACTTGTC GCTTTAGGAT CAACTTTACC 3360

TGTACCTTTT CTCCTTTCCT CCCTTGCCAC CTCAGGTGCA AATCTGAACT CAGTGTCTGC 3420

TTCTTCCATT TTCTCGTCTC TCTCCCCTCT TCCCCCATTA TCCATATGAC ATTATTTTAC 3480

TTCAAATGAC AGCATCAATC TTAAAAAGAT ATACATTAAA ACTAAGGAGT TTTTTTAAAG 3540

AAAGCCTGAA TAAGTTCCTT TCCCTGGTAA CTTTGAAAAG CAGTCAGAGT TGCTATATAG 3600

ATATATGTGG CTCCTTTAAA ATGCTTTGTG TATGTGTGGT GTTTAAAAAA AAAAAAAAAA 3660

TTCGGGGGGG GGCCCGGTNC CCAT 3684

(2) INFORMATION FOR SEQ ID NO: 116:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1965 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 116:

AAGAAAGGGT ATTAAAATTC TAGATCACAT ATGGACCCGG GAAGGTTTTT NACCCTCTGT 60

TAGTGACATC GAGTCTCCCA CTAGACAAAA TAGGTGGAAA AATCTCTCGA GGGCTCACAT 120

TGTTTTGTCA TCTTCAGGAA AAACACCACC AGGCCATACC ACAGCCTGCC CAGTGAGGCG 180

GTCTTTGCCA ACAGCACCGG GATGCTGGTG GTGGCCTTTG GGCTGCTGGT GCTCTACATC 240

CTTCTGGCTT CATCTTGGAA GCGCCCAGAG CCGGGGATCC TGACCGACAG ACAGCCCCTG 300

CTGCATGATG GGGAGTGAAG CAGCAGGAAG GGGCTCCCAA GAGCTCCTGG TGGTGCAGCC 360

TGTGCTCCCC TCAGAAGCTC TGCTCTTCCC AGGGCTCCCG GCTGGTTTCA GCAGGCGACT 420

TTCTTCCAAT GCTGGGCCCA GACTTCTTGC CTGGGTGCTG GCCTGCCCTC TCCGGNCCGC 480

TTGCTGCCTG TCTGCTTTCC TTGGTGGYTT TGCTGGGTGC TGGGCCTGCC CTCTCCGGCC 540

GCTTGCTGCC TGTCTGCTTT CCTTGGTGGC TTTGCTGGGT GCTGGGCCTG CCTTCTCTGG 600

CTGCTTGCTG CCTGTCTGCT TTCCTTGGTG GCTTTGGCTT CTGCACTCCT TGGCGTCASC 660

TCTCAGGTCC TCCATTCACA CGAGGTCCTC CTCGCTCTGG CCGCTCTTGC TGCTCCTGTC 720

TGAAGAWATC AGACTGATTT CCTCTTAAGA CTCCTAGGGA TGTGGTGAAG AGCTGGGACT 780

CAAGTGCAGT CCACGGTGTG AAACATGAGG GARGTGAGGT GTCCGTCCAC TTCCCCCATA 840

AAGGTGTGCA TTTCAGTTAG GCTGCCCCGC CACAGAGCAG GCTTCATCTG CTCTGCCATC 900

CAGCCCCATC TGGATGTGAG GTGGGGTGGA GACATCATGG GGTGATTGCA GAAAGGGGGA 960

GTGGCGGCCC ACGCAGCTTC TGCTGAGGAG CTGACCGCTC TGAGCTGTTC TGTTTCGTAT 1020

TGCTGCTCTG TGTCTGCATG TATTGTGACC GTGCGGCTCC ACCTCTTCCA GCTGCTGCTA 1080

CAGCTGAGGC CTGGATCCCG GCCTTTCCCT GTGACTTACG TGTCTGTCAC CGGCANGCAG 1140

CCCTACAAAT CCTGGTGACC TGCTCTCCCA AGAACAGAGC CTGTCCCCAG ATGTCCCAGT 1200

AGCGATGAGT AACAGAGGTG GCTGTGGACT TCCTCTACTT CTCCTTGCTG GATCAGGGCC 1260

TTCCTGCCTC CCGCTGGGCA GGTCTGGCCT TGCTCTCTTG GCAGGGCCCC AGCCCCTCTG 1320

ACCACTCTGC AGCTCACCAT GCAGCTGATG CCAAAGTTGT GGTGTCCAGT GTGCAGCAGC 1380

CCTGGGAGCC ACTGCCACCT TCAGAGGGGT TCCTTGCTGA GACCCACATT GCTTCACCTG 1440

GCCCCACCAT GGCTGCTTGC CTGGCCCAAC CTAGCGTTCT GTGCCATGCT AGAGCTTGAG 1500

CTGTTGCTCT TCTTCAGGGG AGGAAATAGG GTGGAGAGCG GGAAGGGTCT TGCTCCTAAG 1560

TGTTGCTGCT GTGGCTTTTT TGCCTTCTCC AAAGACGCAC TGCCAGGTCC CAAGCTTCAG 1620

ACTGCTGTGC TTAGTAAGCA AGTGAGAAGC CTGGGGTTTG GAGCCCACCT ACTCTCTGGC 1680

AGCATCAGCA TCCTACTCCT GGCAACATCA GGCCAACGTC CACCCCAGCC TCACATTGCC 1740

AGATGTTGGC AGAAGGGCTA ATATTGACCG TCTTGACTGG CTGGAGCCTT CAAAGCCACT 1800

GGGATGTCCT CCAGGCACCT GGGTCCCATG ACCAGCTCCC CGTCTCCATA GGGGTAGGCA 1860

TTTCACTGGT TTATGAAGCT CGAGTTTCAT TAAATATGTT AAGAATCAAA GCTGTCTTTG 1920

TTCAGGCTGC TATAACAAAA ATATAATAGC CTGGGTGGCT TAAAC 1965

(2) INFORMATION FOR SEQ ID NO: 117:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 503 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 117:

AGTGATCCCC TTGCCTCGGC CTCCCAAAAT GCTGGAATTG TAAGCGTGGG CCTCTGCACC 60

CGGCCTGGTC CGCAATTTAA AAACGCACAG CCACCATTCC CTYTCCAGAA AGCACCCAGA 120

TGCCTTTGGG AGAACCAGCC TCCTCCATGG AGGAAAGCTT GGGATCTGCC TTCCCACCTG 180

GGGAGGAGAG GGATCTGTGG AAAATCCTTC TGACGGACTT CCCCTCAGTG CCTGATCCAT 240

ACTCAATAGT AGAAAAAGTA AGAAATATAC AAAGATAGCA GATACACGGA GACAGTTCCC 300

CAAATAGCTG AGCGAWTAGC GCAGAAGCAA TATTGAAGAC CTAATAGCTG AGACATTTCC 360

AGAACTGATA AAGTGCATCC AGCCACAGAT CAAGCAGCCC AGAAAATTCC AGGCAGCATC 420

AACAAATAAA TAGCCCCACA TGCACCCGTG AAAATGCAGA AGACCAAACA AAAAAGTCCG 480

GTCAACAGCC AGAGTTAAAG AGG 503

(2) INFORMATION FOR SEQ ID NO: 118:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1133 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 118:

GGCACAGCTT GGAATGAACC CCTGTGGATA AGGGGGACTA TTAGATAGAA TAAACATCAA 60

TAAATGCTTG ATGAATAAAC GCTAATCCTA CCTTCCCAGC CTGACACCTC CCAGTGGACA 120

CCACACTTCA CTTGAAGCCT TAGAAACCTT TCCCACCCAT GCTTCCAGCC CTGGCTTCAT 180

GTTGCCATTT CTCACCCCCA GAACAGGCCG CCCGCCTGAA GAAACTACAA GAGCAAGAGA 240

AACAACAGAA AGTGGAGTTT CGTAAAAGGA TGGAGAAGGA GGTGTCAGAT TTCATTCAAG 300

ACAGTGGGCA GATCAAGAAA AAGTTTCAGC CAATGAACAA GATCGAGAGG AGCATACTAC 360

ATGATGTGGT GGAAGTGGCT GGCCTGACAT CCTTCTCCTT TGGGGAAGAT GATGACTGTC 420

GCTATGTCAT GATCTTCAAA AAGGAGTTTG CACCCTCAGA TGAAGAGCTA GACTCTTACC 480

GTCGTGGAGA GGAATGGGAC CCCCAGAAGG CTGAGGAGAA GCGGAACNTG AAGGAGCTGG 540

CCCAGAGGCA ANGAGGAGGA GGCAGCCCAG CAGGGGCCTG TGGTGGTGAG CCCTGCCAGC 600

GACTACAAGG ACAAGTACAG CCACCTCATC GGCAAGGGAG CAGCCAAAGA CGCAGCCCAC 660

ATGCTACAGG CCAATAAGAC CTACGGCTGT KTGCCCGTGG CCAATAAGAG GGACACACGC 720

TCCATTGAAG AGGCTATGAA TGAGATCAGA GCCAAGAAGC GTCTGCGGCA GAGTGGGGAA 780

GAGTTGCCGC CAACCTCCTA GGCGCCCCGC CCAGCTCCCT TTGACCCCTG GGGCAGGGCA 840

GGGGGCAGGG AGAGACAAGG CTGCTGCTAT TAGAGCCCAT CCTGGAGCCC CACCTCTGAA 900

CCACCTCCTA CCAGCTGTCC CTCAGGCTGG GGGAAAACAG GTGTTTGATT TGTCACCGTT 960

GGAGCTTGGA TATGTGCGTG GCATGTGTGT GTGTGTGTGA GAGTGTGAAT GCACAGGTGG 1020

GTATTTAATC TGTATTATTC CCCGTTCTTG GAATTTTCTT CCCATGGGGC TGGGGTACTT 1080

TACATTCAAT AAATACTGTT TAACCCAAAA AAAAAAAAAA AAAAGAAAGA AGN 1133

(2) INFORMATION FOR SEQ ID NO: 119:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1101 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 119:

GGGCACAGCT GAAGCTGCAG ACCTCCCCAG GGGATGGCTC CTCTCCCCCA GGAGCCCCGA 60

GGCAGGGGAG GCAGAAAGCC TGGGCTCTGG GGGGTGGCCT GCGGACAGCT GTGCTGTGGG 120

CCGGGGGCTG GGCCTGTCCC ACAGGGNCGT GGAGCTCGTG GTTCTGAGCA GCCAGCTGGG 180

TGGTGTCTGG GGATAGCTGG GAGGCACAGC GGCTGCCATG TGGGACTGGG ACTGGAGTGC 240

TCCCTGGTCT TGGCCTCTGT GGCTCAGCCT TGCTCTGGTC TGCCTGAGTG CAGGGGCCAA 300

GGGGCACAGG GCCAGTGAGG CCGGCCACGC TCGGGCCCTC ACCTGTGAGA TGGGGTCGGA 360

ATTTKACACA GCCTANGGCT TGGTTCTTGG TKGTNGAMCG TGGACTYCTK AGAACGGGAG 420

TGCTGGTCCT GAAAGGCGTG GTTGGAGACC AGCTGCTTTT CTCGCTGTTT TTCTCTTAGG 480

AGATTAAACA AAAACAGAAA GCACAAGACG AACTCAGTAG CAGACCCCAG ACTCTCCCCT 540

TGCCAGACGT GGTTCCAGAC GGGGAGACGC ACCTCGTCCA GAACGGGATT CAGCTGCTCA 600

ACGGGCATGC GCCGGGGGCC GTCCCAAACC TCGCAGGGCT CCAGCAGGCC AACCGGCACC 660

ACGGACTCCT GGGTGGCGCC CTGGCGAACT TGTTTGTGAT AGTTGGGTTT GCAGCCTTTG 720

CTTACACGGT CAAGTACGTG CTGAGGAGCA TCGCGCAGGA GTGAGGCCCA GGCGCCGAGA 780

CCCAAGGCGC CACTGAGGGC ACCGCGCACC AGAGCGTGAC CTCGGCAGGC TGGACACACT 840

GCCCAGCACA GGCAGACCCA CCAGGCTCCT AGGTTTAGCT TTTAAAAACC TGAAAGGGGA 900

AGCAAAAACC AAAATGTGTG ACTGGGCTTT GGAGGAGACT GGAGCCTCAG CCCTGTCCTG 960

GCCACGGGCC GCTGGGGCTG GTGTGGGTGG GCCTTGTGTG CTGGATTTGT AGCTTATCTT 1020

CCGTGTTGTC TTTGGACCTG TTTTAGTAAA CCCGTTTTTC ATTTTAAAAA AAAAAAAAAA 1080

AAACTTTGGG GGGGGGCCCC N 1101

(2) INFORMATION FOR SEQ ID NO: 120:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 282 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 120:

AGCTTCTCTG TCCAGTCTTG AACTCTGGGS TCTCTTGGAA CTTTCCTCAC CCCTCTCAGC 60

CTGAATATTC CTTCCATGGA TTCCACTCAA CCAGACTTTG GATCTGTGCC TACTTAATCA 120

ACCTTATCTT TGCAATATGT TCGGGCCCAC CTTCCACTCC TTGGTTCTTG TTCCTCCTTG 180

GCCTAACTTG TCCCTTCTCC ACTTCACATC CCCGGTGGGA CAGCATTCCT CCTTCCTCCC 240

AACCTCCCTC CGTCTCARAA AAAAAAAAAA AAAAAAAAAA TT 282

(2) INFORMATION FOR SEQ ID NO: 121:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2635 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 121:

TAAGGGGGTG TGTGCTCACC TCCTCCTGAC CCTTAACACT CCTGTCCTGC CCAGACCAAC 60

AGAGAGAGCT GTCCCTGAGA CCCCGGAGAG AAGCAGCTGC CGAAAGCTGC AGCCTTTCCG 120

CACTCTGAGA CCATGATCTT CCTCCTGCCA GGGGAGAGCC ACCCACAGGC CATGTCCAGC 180

CCCACTTCCC TCAGCCCCCA GGGYTTCCTT CTGGCCCCTC TGAGGATTCC CTAGGGCTGC 240

CCCGCAGAGG GGYTTCCCCA AGCTCTGTTT TGAAGCCTGC AATGTGGAAA AGTGAGAAGT 300

CAGAGGGAAC AGGACAGGTG CAGCCGGGCT CTGAGGCCAC ACCTCACACC TCGCTGTTCC 360

CCAACATCCC CTGAGCAGTG TGAGCTCATC TCACCAGATG AGAAGAGGCC CTGTGCATTT 420

YTTTTGTTTG TTTGTTGCTG TTTTCCCCCA CCCATCCAGT TCTCCTCAGC AAAGCAAATT 480

CCTTAACACC TTTGGTGGAG AATTTCTTAC CCAGACTTGG GGCTGTGATG CCCTTCAGTG 540

CGTGGTGAGT GCAGCGTGTG TGCGTGTGCC TGTGTGTGAA CCTGGGGGCC ATCCTGGTGG 600

CCTGGGAGCG TGAGGAGAGG CCCCCTGTGT GCTGGGTGAG TGGTGGGTGT GGGGTCAATG 660

CAGTGAGGCT CTCTGGGTGA GGCTCCCAAC CTGGCAGTCC CCAGCCTCCC AGCATCTGTG 720

AGCGTCTGTT GGACTTTACA GAAGAGCCTC ATCCYGTCTG CCCCTCACTC TGCCCTGGAA 780

TCAACATCTT CCGAGTCCTT CTTGGGGGAA ATAGCAGAGC CCCACTTAAC TCCATAAACT 840

GCTTCCCATT CCGCAGCCCA GTTCTGATTG TTGAGGTGTC GCGTCGTTCC AGGTCCCCCA 900

GTCCCCTCTT TCTCCTGTCC TCTCTCTGTC CTTCACCTCC CCACTCCAGC CCCGGCTCAG 960

TTCAGGGAAA TGCTGTTCCA YATCAGCCCT CTGCTCTCTG AGGCAGCCGC GCCTCTGACT 1020

CGGAGCTACT TGAAACTTCT GCTCTTGCTA GGATTGGAGT CTACCTATCT CTTCCATTTG 1080

TCCCAGCTGG AGTTCTGGAA CTTTCCTCCT CGGGGTGGGG GTGGGGGTTG TTAAGGATGC 1140

TGGGGGGCCT GGGGAAGGAA GGAGTTCAGA GGAAGGGTGT CCCCTGTCCT CTTGATGTCA 1200

CCCTCCGCTC CTGGGACACG TGCTCTCTCT GTCTCTGGGT CTTCTGGGTG TGCACGTTTG 1260

TGTGTCCTTG TAAATATGTT TTAGGAAGAA AGCAAAAGGG ACTGAACTAG CCTCTGGTAG 1320

GATTGCAGGG GTCCAGCCTT GCCTGTTTCC GAAGCCCCCA CACTGCCTTT CGCCCCACTG 1380

AGACTGGTCC CCTCAAAAGG TAGACAAAAC AGCAGCTCCC TGTGGAGCTG AAGGGCGGCC 1440

TCAAAGTGGC TTTTTGTTAG ACAAGGTTAA GGTTTCCTCA TGAGCAAGGT TGTAGATCGG 1500

TCCTTCCTCA GCTCCTTGAT TTGTGACCTT GACCAAGGGG CCTGCCACCC AGCCCCTCCA 1560

GTGCCCTCTC CTCGATGCCT CGCTCCTTCC TGCCCCCACT CCCCTGGCTT AGGCAGGTAG 1620

GGGAATTAGG GCCATGCTGG AAGAAGCTTA ACCATGTGTT CAAAGAACGG TTTCTTGCTT 1680

GCTTGGTCCT GGAACTCCCC TTGGCTGCCC CAGGCCTCCT TGGCCCAACG GTGCTGGGGG 1740

AGGTGGATGT CAGATCTGGT AGGTTGCAGC AGAGAAAATA AATGTGCCTT GAGAGACCAC 1800

TCAGAGAGGG TCCAAGGGTG ATGGAGAAGG AAGCATGGCC TGGGAGCTTG GAAGGGARGG 1860

GTGGTGGGTG GCGGCATCTT GACTGCCCCC TGTTGTCCCA CACGTGGGGG GTGGTCACCC 1920

CYCTTCACTC CAGCCCGCCT GCCTTCAGCC TTCCATGAGC TTCACCTGCT TCCAACTTCA 1980

CTTTGGAGGG GGTGGGGTCC GTTGGCATCA ACACGGGGAC CCTCTGCTTC ACCAAAGCCC 2040

GAGCCCTCAG CCCCTGGGGA GAACAAATGG CTGAGCTTTG ATACCTGGGG TCGTCGAGAG 2100

GCTGCGGGCT GGCGGCAGTC CCAGGGGAGA GACACCACAG AAGGAGACCC AGACATCCCG 2160

AGGAAGTTCC CAGCAGAGCA AACTGCTTTC CAGCCTGAAG CCTGCTTAAA CTGTGTGATG 2220

TGCAATAACT GAGCTTAGAG TTAGGAATTG TGTTCAAGTG CTTGGATTTC CGTCTGTAGA 2280

TTTAACTGCT GAAATTGTAT CTCTCAGTAA TTTTAGATGT CTTTTAAAAA ATTGAAAAAC 2340

AAAGTGTTAG ACTGTGTGCG TGTGCGTTGA TGGGCACTCA AGAGTCCCGT GAGTCATCCA 2400

GCCCTGCCTT TCCCCTGCGC CCCCATCCTC TCACGTCCCG CCCAGCCTCC ACTTGGGGAC 2460

CCTGCCTCGT GTCGTCTTTA TCTGCCTATT ACTCAGCCTA AGGAAACAAG TACACTCCAC 2520

ACATGCATAA AGGAAATCAA ATGTTATTTT TAAGAAAATG GAAAATAAAA ACTTTATAAA 2580

CACCAAAAAA AAAAAAAAAA ACCCNGGGGG GGGGCCGGTA ACCCATTTCG CCTAA 2635

(2) INFORMATION FOR SEQ ID NO: 122:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 994 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 122:

GAATTCGGCA GAGGTTCGGC GAAGATAGGG AATAAGGAAG CACAGGAGTA GGGGAGAAGG 60

AAGCACAGGA GTAGGGGAGA TATACAGCGG TCAGGATAAG GGGGAAAGGG CGGTGGTTGC 120

SCAAGAGGTG AAACAAGATG TGAGAGACAA GGGGTAGGGA AGAAATGGGG CAGCGGTTAG 180

GTTCAGAAGC GCATAGACCG TGGCGGACGG GCAATGCGAG GGGCACAGAA AGGAACTGAG 240

GGGTGGGCTA TTTTAARGGA GATGGTCCTT CAGCCCTCTT YTTTTCTGCG TAGTTCTCCT 300

CCTCCAGGCC GCGCGCGGAT ATGTCGTCCG GAAACCAGCC CAGTCTAGGC TGGATGATGA 360

CCCACCTCCT TCTACGCTGC TCAAAGACTA CCAGAATGTC CCTGGAATTG AGAAGGTTGA 420

TGATGTCGTG AAAAGACTCT TGTCTTTGGA AATGGCCAAC AAGAAGGAGA TGCTAAAAAT 480

CAAGCAAGAA CAGTTTATGA AGAAGATTGT TGCAAACCCA GAGGACACCA GATCCCTGGA 540

GGCTCGAATT ATTGCCTTGT CTGTCAAGAT CCGCAGTTAT GAAGAACACT TGGAGAAACA 600

TCGAAAGGAC AAAGCCCACA AACGCTATCT GCTAATGAGC ATTGACCAGA GGAAAAAGAT 660

GCTCAAAAAC CTCCGTAACA CCAACTATGA TGTCTTTGAG AAGATATGCT GGGGGCTGGG 720

AATTGAGTAC ACCTTCCCCC CTCTGTATTA CCGAAGAGCC CACCGCCGAT TCGTGACCAA 780

GAAGGCTCTG TGCATTCGGG TTTTCCAGGA GACTCAAAAG CTGAAGAAGC GAAGAAGAGC 840

CTTAAAGGCT GCAGCAGCAG CCCAAAAACA AGCAAAGCGG AGGAACCCAG ACAGCCCTGC 900

CAAAGCCATA CCAAAGACAC TCAAAGACAG CCAATAAATT CTGTTCAATC ATTTAAAAAA 960

AAAAAAAAAA AAAAAAAAAA AAAAAGGGGA GGGG 994

(2) INFORMATION FOR SEQ ID NO: 123:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1542 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123:

GGCASAGCCA CCTCGGCCCC GGGCTCCGAA GCGGCTCGGG GGCGCCCTTT CGGTCAACAT 60

CGTAGTCCAC CCCCTCCCCA TCCCCAGCCC CCGGGGATTC AGGCTCGCCA GCGCCCAGCC 120

AGGGAGCCGG CCGGGAAGCG CGATGGGGGC CCCAGCCGCC TCGCTCCTGC TCCTGCTCCT 180

GCTGTTCGCC TGCTGCTGGG CGCCCGGCGG GGCCAACCTC TCCCAGGACG ACAGCCAGCC 240

CTGGACATCT GATGAAACAG TGGTGGCTGG TGGCACCGTG GTGCTCAAGT GCCAAGTGAA 300

AGATCACGAG GACTCATCCC TGCAATGGTC TTAACCCTGC TCAGCAGACT CTCTACTTTG 360

GGGAGAAGAG AGCCCTTCGA GATAATCGAA TTCAGCTGGT TAMCTCTACG CCCCACGAGC 420

TCAGCATCAG CATCAGCAAT GTGGCCCTGG CAGACGAGGG CGAGTACACC TGCTCAATCT 480

TCACTATGCC TGTGCGAACT GCCAAGTCCC TCGTCACTGT GCTAGGAATT CCACAGAAGC 540

CCATCATCAC TGGTTATAAA TCTTCATTAC GGGAAAAAGA CACAGCCACC CTAAACTGTC 600

AGTCTTCTGG GAGCAAGCCT GCAGCCCGGC TCACCTGGAG AAAGGGTGAC CAAGAACTCC 660

ACGGAGAACC AACCCGCATA CAGGAAGATC CCAATGGTAA AACCTTCACT GTCAGCAGCT 720

CGGTGACATT CCAGGTTACC CGGGAGGATG ATGGGGCGAG CATCGTGTGC TCTGTGAACC 780

ATGAATCTCT AAAGGGAGCT GACAGATCCA CCTCTCAACG CATTGAAGTT TTATACACAC 840

CAACTGCGAT GATTAGGCCA GACCCTCCCC ATCCTCGTGA GGGCCAGAAG CTGTTGCTAC 900

ACTGTGAGGG TCGCGGCAAT CCAGTCCCCC AGCAGTACCT ATGGGAGAAG GAGGGCAGTG 960

TGCCACCCCT GAAGATGACC CAGGAGAGTG CCCTGATCTT CCCTTTCCTC AACAAGAGTG 1020

ACAGTGGCAC CTACGGCTGC ACAGCCACCA GCAACATGGG CAGCTACAAG GCCTACTACA 1080

CCCTCAATGT TAATGACCCC AGTCCGGTGC CCTCCTCCTC CAGCACCTAC CACGCCATCA 1140

TCGGTGGGAT CGTGGCTTTC ATTGTCTTCC TGCTGCTCAT CATGCTCATC TTCCTTGGCC 1200

ACTACTTGAT CCGGCACAAA GGAACCTACC TGACACATGA GGCAAAAGGC TCCGACGATG 1260

CTCCAGACGC GGACACGGCC ATCATCAATG CAGAAGGCGG GCAGTCAGGA GGGGACGACA 1320

AGAAGGAATA TTTCATCTAG AGGCGCCTGC CCACTTCCTG CGCCCCCCAG GGCCCTGTGG 1380

GGACTTGCTG GGGCCGTCAC CAACCCGGAC TTGTACAGAG CAACCGCAGG GGCCGSCCCT 1440

CCCGNTTGTT CCCCAGCCCA CCCACCCCCT TGTTACAGAA TGTYTKGTTT GGGGTGCGGT 1500

TTTGTWATTG GTTTNGGATN GGGGAAGGGA GGGANGGCGG GG 1542

(2) INFORMATION FOR SEQ ID NO: 124:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1390 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 124:

CAAGCTCTAA TACGACTCAC TATAGGGAAA GCTGGTACGC CTGCAGGTAC CGGTCCGGAA 60

TTCCCGGGTC GACCCACGCG TCCGGGCCTC AGGGTGGACG CATGGTTCTG CACTGAGGCC 120

CTCGTCATGG TGGCGCCTGT GTGGTACTTG GTAGCGGCGG CTCTGCTAGT CGGCTTTATC 180

CTCTTCCTGA CTCGCAGCCG GGGCCGGGCG GCATCAGCCG GCCAAGAGCC ACTGCACAAT 240

GAGGAGCTGG CAGGAGCAGG CCGGGTGGCC CAGCCTGGGC CCCTGGAGCC TGAGGAGCCG 300

AGAGCTGGAG GCAGGCCTCG GCGCCGGAGG GACCTGGGCA GCCGCCTACA GGCCCAGCGT 360

CGAGCCCAGC GGGTGGCCTG GGCAGAAGCA GATGAGAACG AGGAGGAAGC TGTCATCCTA 420

GCCCAGGAGG AGGAAGGTGT CGAGAAGCCA GCGGAAAYTC ACCTGTCGGG GAAAATTGGA 480

GCTAAGAAAC TGCGGAANNT GGAGGAGAAA CAAGCGCGAA AGGCCCAGCK TGAGGCAGAG 540

GAGGCTGAAC GTGARGWGCG GAAACGACTC GAGTCCCAGC GCGAATGAGT GGAAGAAGGA 600

GGAGGAGCGG CTTCGCCTGG AGGAGGAGCA GAAGGAGGAG GAGGAGAGGA AGGCCCGCGA 660

GGAGCAGGCC CAGCGGGAGC ATGAGGAGTA CCTGAAACTG AAGGAGGCCT TTGTGGTGGA 720

GGAGGAAGGC GTAGGAGAGA CCATGACTGA GGAACAGTCC CAGAGCTTCC TGACAGAGTT 780

CATCAACTAC ATCAAGCAGT CCAAGGTTGT GCTCTTGGAA GACCTGGCTT CCCAGGTGGG 840

CCTACGCACT CAGGACACCA TAAATCGCAT CCAGGACCTG CTGGCTGAGG GGACTATAAC 900

AGGTGTGATT GACGACCGGG GCAAGTTCAT CTACATAACC CCAGAGGAAC TGGCCGCCGT 960

GGCCAACTTC ATCCGACAGC GGGGCCGGGT GTCCATCGCC GAGCTTGCCC AAGCCAGCAA 1020

CTCCCTCATC GCCTGGGGCC GGGAGTCCCC TGCCCAAGCC CCAGCCTGAC CCCAGTCCTT 1080

CCCTCTTGGA CTCAGAGTTG GTGTGGCCTA CCTGGCTATA CATCTTCATC CCTCCCCACC 1140

ATCCTGGGGA AGTGATGGTG TGGCCAGGCA GTTATAGATT AAAGGCCTGT GAGTACTGCT 1200

GAGCTTGGTG TGGCTTGGTG TGGCAGAAGG CCTGGCCTAG GATCCTAGAT AAGCAGGTGA 1260

AATTTAGGCT TCAGAATATA TCCGAGAGGT GGGGAGGGTC CCTTGGAAGC TGGTGAAGTC 1320

CTGTTCTTAT TATGAATCCA TTCATTCAAG AAAATAGCCT GTTGCAAAAA AAAAAAAAAA 1380

AAAAACTCGA 1390

(2) INFORMATION FOR SEQ ID NO: 125:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1288 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 125:

GGCGCGCGGG TGAAAGGCGC ATTGATGCAG CCTGCGGCGG CCTCGGAGCG CGGCGGASCA 60

GACGCTGACC ACGTTCCTCT CCTCGGTCTC CTCCGCCTCC AGCTCCGCGC TGCCCGGCAG 120

CCGGGAGCCA TGCGACCCCA GGGCCCCGCC GCCTCCCCGC AGCGGCTCCG CGGCCTCCTG 180

CTGCTCCTGC TGCTGCAGCT GCCCGCGCCG TCGAGCGCCT CTGAGATCCC CAAGGGGAAG 240

CAAAAGGCGC ATCCGGCAGA GGGAGGTGGT GGACCTGTAT AATGGAATGT GCTTACAAGG 300

GCCAGCAGGA GTGCCTGGTC GAGACGGGAG CCCTGGGGCC AATGGCATTC CGGGTACACC 360

TGGGATCCCA GGTCGGGATG GATTCAAAGG AGAAAAGGGG GAATGTCTGA GGGAAAGCTT 420

TGAGGAGTCC TGGACACCCA ACTACAAGCA GTGTTCATGG AGTTCATTGA ATTATGGCAT 480

AGATCTTGGG AAAATTGCGG AGTGTACATT TACAAAGATG CGTTCAAATA GTGCTCTAAG 540

AGTTTTGTTC AGTGGCTCAC TTCGGCTAAA ATGCAGAAAT GCATGCTGTC AGCGTTGGTA 600

TTTCACATTC AATGGAGCTG AATGTTCAGG ACCTCTTCCC ATTGAAGCTA TAATTTATTT 660

GGACCAAGGA AGCCCTGAAA TGAATTCAAC AATTAATATT CATCGCACTT CTTCTGTGGA 720

AGGACTTTGT GAAGGAATTG GTGCTGGATT AGTGGATGTT GCTATCTGGG TTGGCACTTG 780

TTCAGATTAC CCAAAAGGAG ATGCTTCTAC TGGATGGAAT TCAGTTTCTC GCATCATTAT 840

TGAAGAACTA CCAAAATAAA TGCTTTAATT TTCATTTGCT ACCTCTTTTT TTATTATGCC 900

TTGGAATGGT TCACTTAAAT GACATTTTAA ATAAGTTTAT GTATACATCT GAATGAAAAG 960

CAAAGCTAAA TATGTTTACA GACCAAAGTG TGATTTCACA TGTTTTTAAA TCTAGCATTA 1020

TTCATTTTGC TTCAATCAAA AGTGGTTTCA ATATTTTTTT TAGTTGGTTA GAATACTTTC 1080

TTCATAGTCA CATTCTCTCA ACCTATAATT TGGGAATATT GTTGTGGTCT TTTGTTTTTT 1140

CTCTTAGTAT AGCATTTTTA AAAAAATATA AAAGCTACCA ATCTTTGTAC AATTTGTAAA 1200

TGTTAAGAAT TTTTTTTATA TCTGTTAAAT AAAAATTATT TCCMACAACC TTAAAAAAAA 1260

AAAAAAAAAA AAAAAAAAAA AAAAANAA 1288

(2) INFORMATION FOR SEQ ID NO: 126:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1517 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 126:

AGTGGCTTAA AGGCATCGTT TTAGGGATTA CTGGGAAGTA TCTTCAAAGT AATACATGAG 60

AAACATTCCT TCCTAAATCC TTTATTATAT TGAATATCGT ATTAATTGGT TTTCAGAGGT 120

TAAATTAACC ATGTATTCCT GCAATAAATG TCACTTGTNT CTTGTATATA ATCTTTTTTA 180

TATATTACCG GATTGATTCA TTAGTATTTT GTTGAGGATT TTTGTGTCTA TATTCATAAG 240

AGATGCTGGT CTGCAGTTTT CTTTTTTTGT GATAATCTGG TTTTTGTATC AGTAATACAG 300

GCCCCATGAA ACGAGTTGGG AAGTGTTCAC CTCTCTTGTA TTTTTTCAAG AGTTTGTGAA 360

GAATTGCTAT TAATTCTTTA AATGTTTGGT AGAATCTACC ATTGAAATCA TGTGTCCTGG 420

GCTTTTTTTT GAGGGAAGTG TTCTGATAAC TAATTCAGTA TCTACTTTTT ATAGCTCTGT 480

TCAGATTTTG CTTCTTCCTG AGTTAGTTTT GGTAATTTGT GTATCTCTAG GARTTTGTCC 540

ATTTCATTTA TCTCATTTGT TGGCATAAAT TAAACTAAAT TTGGCCTGAG CCTACCTGTA 600

TATCTTGAGT CCCTCTGTAA GGAACTGTAG CCTAACTTGT ACATAAACAA ACTGAAATCC 660

TAAATTAGGA ATGTAGTTTT TGTAACAGCT CCTGAGTCTC AGGCAGTCAC AGCAGYCAAG 720

TCTGTCAATT GCAGGCTGCT AACTAAGCAG CCCATGSTCA AATGAGGCAA AAACCTTTGC 780

TTTTAACACA TAGTATAGCT TTGTAATCCT TTTCTTGCAC ACTCGGGTAA TTTCTTCCTT 840

TTTCATTCCC KGWATTTTCC AKGAATATGA RTCTYCCTTT TTTCCCCTCC TGTCAGTCTA 900

GCTAATGGTT TGTCAATTTT GTTGATCTTT TGAARAACAA ACCTTTGGTT CCACTTTCTT 960

GTTGCATATG CTGARTATTC TCATAATTGG AGTGGAAAGC TGATCTTTGA TTACTTATTT 1020

TACTTAGGGC TGAGGAGTTC ATGGACTTCG CAAAACCTCC TTGAATCTAA ATTGCATCTT 1080

CTTTCCTGGT TTCTGGGCTG AAACATGTTT TTTCCCATCT WANAWACCCT TGGTCTTTTC 1140

ATKGGCGATT AAGACTAGAG AAAGTTCTAG ATMCCTTGTC CTTTTATGCT GTCATTTTGT 1200

TTAAAGGCTT TCTATGTAGT AAAACTATCT ATATAGACAA AATAGAGCCT TGAGTTGTGG 1260

TCTTGAATTT GATCAACATG ATTTACCACA TTCTGTACTG GATATTTCTT CACCTGCTGC 1320

TACTGTAAAC CATTTTATTC TTGGATCTTC TGTAGAGTAT ATTATCACAG GTACTTTTTA 1380

CAGGGGTGTC TAATCTTTTG GCTTCCCTGG GCACATTGAA AGAAGAAGAA TTGTCTTGGG 1440

CCACACATCA AATACGCTAA CACTAATAAT AGTTGATGAG CTAAAAAAAA AAAAAAAAAG 1500

GCAAAAAAGN CCCAAAA 1517

(2) INFORMATION FOR SEQ ID NO: 127:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1073 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 127:

TGAATCTATT CTTTGAACAT TCTACAACAA GAATTACATT ATACTGTTAT ACCAGAGTAC 60

TTCTGCAGTG TGAAATAGAT TGGTTTGGAA AATGAACCTG GCTTTGCTAT AAATTACATT 120

CACAGGCCTT TTTGCAAATG TGTAACTTGC CTATCAAAGT AGTTTGTAGG GCAAATGCAG 180

AATATATGTC TCCATCTGGT AAAGTACCTT WTAYTCATGT GGGAAATCAA GTAGTATCAG 240

AACTTGGTCC AATAGTCCAA TTTGTTAAAG CCAAGGGCCA TTCTCTTAGT GATGGGCTGG 300

AGGAAGTCCA AAAAGCAGAA ATGAAAGCTT ACATGGAATT AGTCAACAAT ATGCTGTTGA 360

CTGCAGAGCT GTATCTTCAG TGGTGTGATG AAGCTACAGT AGGGRMGATC ACTCATGMTA 420

GGTATGGWTC TCCTTACCCT TGGCCTCTGW WTCATATTTT GGCCTATCAA AAACAGTGGG 480

AAGTCAAACG TAAGNTGAAA GCTATTGGAT GGGGAAAGAA GACTCTGGAC CAGGTCTTAG 540

AGGATGTAGA CCAGTGCTGT CAAGCTCTCT CTCAAAGACT GGGAACACAA CCGTATTTCT 600

TCAATAAGCA GCCTACTGAA CTTGACGCAC TGGTATTTGG CCATCTATAC ACCATTCTTA 660

CCACACAATT GACAAATGAT GAACTTTCTG AGAAGGTGAA AAACTATAGC AACCTCCTTG 720

CTTTCTGTAG GAGAATTGAA CAGCACTATT TTGAAGATCG TGGTAAAGGC AGGCTGTCAT 780

AGAGTTATGT GTTAGTCTCA GGAGTCTTAA CTTTTGAAAT ATGTTTTACT TGAATGTTAC 840

ATTAGATATT GGTGTCAGAA TTTTAAAACC AAATTACTGC TTTTTGAAAC CTCAAATTAT 900

ATAATGTATC TTATGTATGT GCTTTATATT GTTATTTGTG TATACATTAA AATAATTCTG 960

AATTATTTAA TCTGATATGT TGTATTCTGT ATCTTGAAAT TTTTGTTTCC TTGAAACATG 1020

CATGCATTTA AAAATAAAGC TTAAACAACT GTAAAAAAAA AAAAAAAAAA CTC 1073

(2) INFORMATION FOR SEQ ID NO: 128:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 300 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 128:

CAACCCCTGC CTTTTTTTTG TTTTCCATTT GCTTGGTAGA TCTTCCTCCA TCCCTTTATT 60

TTGAGCCTAT GTGTGTCTCT GCCCGTGAGA TGAGTCTCCT GAATACAGCA CACTTACTGG 120

TCTTGACTCT GTATCCAATT TGCCAGTCTG TGTCTTTCAT TTGGAGCATT TAGCCCATTT 180

ACATTTAAGG TKAATATTGT TATGTGTGAA TTTRATCYTR TCATTATGWT GTTAGCTGGT 240

TATTTTGCTT GTTAGTTGAT GCAGTTTCTT CCNGGCATCA ATGGTCTTTA CAANTTGGCA 300

(2) INFORMATION FOR SEQ ID NO: 129:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1275 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 129:

GGCAGAGCCT GTCCCTGCTG CCCCTGCAAA AAAAACCCCC TCTGGTGTGA GCAGGATGGT 60

TGGAGGTTAT GTGAGCTCCT TCTCCTTTCC TCCAGTTTCC TCTTCCCTTC TCCTCCCTGC 120

CTCTTTTGCT TTTCCCTTTC TTCCTGGTAC CCCCTGCCCA TTCCTGTATT TTCTCCCATC 180

GCCATTCTCC CCTCTCCCAC TGTCCCTAAC CCGTTCAAAC TCTTTCCTCT TAAATGGTTG 240

AGATTTTCTC TCACCAAGCA CACCCCAGTA TTAATTAAAC TAGCTGCAAA CAGGCAGCAA 300

GTGGTCTACC ATGACAGATG GGTTTTGTGT GTGTGTGTGT GTGTGTAATT GTAATAAAAC 360

ATATTGARTC ACTCAATAAA CACAGAGTGT CTACTACATG TATCARGCAC TATCATAGAT 420

GCTAATTAAC GAAACTGAAA TGGCCAGGCC CTCACAGTGG CTCATGCCTA TAATCCCAGC 480

ACTTTGGGAG GATGAGGCAG GAGGATCACT TGAGGCCGGG AGTTCAAGAC CAGCCTGGGC 540

AACATAGTAA GACTCCATCT CTACAAAAAA AAAATTTTTT TTATTATACT TTAAGTTTTG 600

GGTTACATGT GCAGAACGTG TAGTTTTGTT ACATAGGTAT ATACGTGCCC TGGTAGTTTG 660

CTGCACCCAT CAACCCATCA CCTACATTAG GTATTTCTCC TAATGTTACC CCTCTCCTAG 720

CCCCCCACCC CGTGACAGGC CCTGGTGTGT GATGTTCCCC TCCCTGTGTC CATGTGTTCT 780

CATTGGTCAA CTCTCACCTA TGGAGTGAGA ACATGTGGTA TTTGGTTTTC TGATCTTGTG 840

ATAGCTTGCT GAGAATGTKG GTTTCCAGCT TTATCCACGT CCCTGCAAAG GGCATAAACT 900

CATCCCTTTT TATGGCTGCA TAGTGTTCCA TGGTGTATAC GTGCCACATT TTCTTAATCT 960

ATCATTGATG GACAAGTTTT GCTATTGTGA ATAGTGCCAC AATAAACATA CGTGTGCGTG 1020

TGTCTTTATA GCAGCATGAT TTATAATCCT TTGGGTATAT ACCCAGTAAT GGGATCACTG 1080

AGTCAAATGG TATTTCTCGT TCTAGATCCG TAAGGAATTG CCACACTGTC TTCCACAATG 1140

TTTGAACTAA TNTACACTCC CACCAACAGT GTAAAAGTGT TTCTATTTTT CCACAACCTC 1200

TCCAACATCT GTTATTTCCT GACTTTTTAA TGAACGTCAT TCTAACTGGC GTGAGATGGT 1260

ATCTCATTGT GGTTT 1275

(2) INFORMATION FOR SEQ ID NO: 130:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 472 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 130:

CNGAAACCCC GTGAACCCTC CCCGGGTTAA AAAGCCCCCC CTAAATGGGG GGAACGCYTC 60

ACACGTTATA AAAAAGCACT AGAATGTTTT GAAAGCGAGA AACAACAGCT GTGTAGGGTA 120

GCTAGCAGTT AGTGTTGTAC AGAAGACAGA TATTTGTGCA TTTYTGCATT TTCTAAGTTT 180

GCTGCAATGA GCATGTATTA CTTTCATAGT TATAAAACAC ATGCAAAATG CCCTTTTAAA 240

ATGAAAAAAA ATCCATGAGT GTAAGTGATA TATATGCTTT GGAAAGCCTG GGACGGTCAT 300

TGTTTACTCT CAATAGTATG TGTTTGCCTT TGTCTTTTTG AGACATTTTG TTTTAATCTG 360

TTGATGACAA TAACCTGTTG ATAATATAAC TTGATAACAA ATAAAATGAC TTATGATTGA 420

AWMAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA NN 472

(2) INFORMATION FOR SEQ ID NO: 131:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1950 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 131:

ACCTCTCAGA ATCTTCTCTC AGCAACCTGA GTCTTCGCCG TTCCTCAGAG CGCCTCAGTG 60

ACACCCCTGG ATCCTTCCAG TCACCTTCCC TGGAAATTCT GCTGTCCAGC TGCTCCCTGT 120

GCCGTGCCTG TNATTCGCTG GTGTATGATG AGGAAATCAT GGCTGGCTGG GCACCTGATG 180

ACTCTAACCT CAACACAACC TGCCCCTTCT GCGCCTGCCC CTTTNTGCCC CTGCTCAGTG 240

TCCAGACCNT TGATTCCCGG CCCAGTGTCC CCAGCCCCAA ATCTGCTGGT GCCAGTGGCA 300

GCAAAGATGC TCCTGTCCCT GGTGGTCCTG GCCCTGTGCT CAGTGACCGA AGCTCTGCCT 360

TGCTCTGGAT GAGCCCCAGC TCTGCAACGG GCACATGGGG GGAGCCTCCC GGCGGGTTGA 420

GAGTGGGGCA TGGGCATACC TGAGCCCCCT GGTGCTGCGT AAGGAGCTGG AGTCGCTGGT 480

AGAGAACGAG GGCAGTGAGG TGCTGGCGTT GCCTGAACTG CCCTCTGCCC ACCCCATCAT 540

CTTCTGGAAC CTTTTGTGGT ATTTCCAACG GCTACGNCTG CCCAGTATTC TACCAGGCCT 600

GGTGCTGGCC TCCTGTGATG GGCCTTCGMA CTCCCAGGCC CCATCTCCTT GGCTAACCCC 660

TGATCCAGCC TCTGTTCAGG TACGGCTGCT GTGGGATGTA CTGACCCCTG ACCCCAATAG 720

CTGCCCACCT CTCTATGTGC TCTGGAGGGT CCACAGCCAG ATCCCCCAGC GGGTGGTATG 780

GCCAGGCCCT GTACCTGCAT CCCTTAGTTT GGCACTGTTG GAGTCAGTGC TGCGCCATGT 840

TGGACTCAAT GAAGTGCACA AGGCTGTGGG GCTCCTGCTG GAAACTCTAG GGCCCCCACC 900

CACTGGCCTG CACCTGCAGA GGGGAATCTA CCGTGAGATA TTATTCCTGA CAATGGCTGC 960

TCTGGGCAAG GACCACGTGG ACATAGTGGC CTTCGATAAG AAGTACAAGT CTGCCTTTAA 1020

CAAGCTGGCC AGCAGCATGG GCAAGGAGGA GCTGAGGCAC CGGCGGGCGC AGATGCCCAC 1080

TCCCAAGGCC ATTGACTGCC GAAAATGTTT TGGAGCACCT CCAGAATGCT AGAGACCTTA 1140

AGCTTCCCTC TCCAGCCTAG GGTGGGGAAG TGAGGAAGAA GGGATTCTAG AGTTAAACTG 1200

CTTCCCTGTT GCCTTCATGG AGTTGGGAAC AGGCTGGGAA GGATGCCCAG TCAAAGGCTC 1260

CAAGCGAGGA CAACAGGAAG AGGGATCCAC TGTTACCAAA AGTCCTGATT CCCCCATCAC 1320

CAACCTACCC AGTTTGTTCG TGCTGATGTT GGGGGAGATC TGGGGGGAGT TGGTACAGCT 1380

CTGTTCTTCC CTTGTCCTAT ACCGGGAACT CCCCTCCAGG GTACCCACAG ATCTGCATTG 1440

CCCTGGTCAT TTTAGAAGTT TTTGTTTTAA AAAACAACTG GAAAGATGCA GAGCTACTGA 1500

GCCTTTGCCC TGAATGGGAG GTAGGGATGT CATTCTCCAC CAATAATGGT CCCTCTTCCC 1560

TGACGTTGCT GAAGGAGCCC AAGGCTCTCC ATGCCTTTCT ACCTAAGTGT TTGTATTTTA 1620

TTTTAAATTA TTTATTCTGG AGCCACAGCC CCCTTGCTTA TGAGGTTCTT ATGGAGAGTG 1680

AGAAAGGGAA GGGAAATAGG GCACCATGGT CCGGTGGTTT GTAGTTCCTT CAAAGTCAGG 1740

CACTGGGAGC TAGAGGAGTC TCAAGCTCCC CTTAGGAAGA ACTGGTGCCC CCTCCAGTCC 1800

TAATTTTTCT TGCCTGCCCC GCCTTGGGGA ATGCCTCACC CACCCAGGTC CTGACCTGTG 1860

CAATAAGGAT TGTTCCCTGC GAAGTTTTGT TGGATGTAAA TATAGTAAAA GCTGCTTCTG 1920

TCTTTTTCAA AANAAAAAAA AAAAAAAACT 1950

(2) INFORMATION FOR SEQ ID NO: 132:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 990 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 132:

TGGAAGATTT AAAATAGGTT TCATATTTCT CTTGAATATG AATATATAAG CTTGAATAAG 60

CTTGAGTCCT TATTATTATG AAATTTTCCT TATTATTTCT ACCAATGCTT CTTATATTAA 120

AGCCTGATCT TTTTCATATT AGTATATGTA CATTAGCTGC CTGTGGATTA ACATTTCCAT 180

GAAATGTATT TTTGCATTGT TTGATCTTAA ACTTTTTGTG TCTTTATATA AGGTATGCTY 240

CTTTTAAGCA TGATATTTTT AACCACAATA GTTGAAAGAC AATCTYCACC TTTTACTTGT 300

ATATTTACAT GTAATGTAAT TTTTGATGCA TATTACGTCT TATTATTTAA CCAACCTATT 360

TTATTTTATC TAGGGCATTT TTCAGAAAGC CTTATTTTCT TGTATTAATC AAATATTTTT 420

AYCATTGTAT TTTCCYCTAT TAGTTAGKAA TACGKTACYC YAAATATATA TTGTGGSTAT 480

TTTCAGAATT GCAATATGCC TCCTTAATTT ATTAGAGGCT AACCTAAATT ATTACTTTTA 540

CCACTTACTT GAAAATTCTG GAACTTTAGA ACATTTATTG TTTTATGCAT TTTAATTCTA 600

CTTGTATTTT TACTACTCCT AAACATTATT ATTGTTTTAG ACAAGCCAAA ATATATNTTG 660

TTATTATCTT ATYCTCCATT TCTTTCTGTA TTTTTATGCC ACTATGTATG CTCAATTTCC 720

TTCTATGTGA TGAACCTAAT TCAGTACTTT TGTTTTTTAA TCTGTGCAGG TAGCCTGGCC 780

ATTAAATTTT TACTTTTGGT TTGCTGAAAA AATTGTGTTT ATTTCTATAT GCATACTTAT 840

GCATATAGAA TNCTAGGTNG ACATATTTTT AGTATTTATA AATGTAAAGT CATTWATTKG 900

GCTTCTATCA TTTCKGTKGA GAAATCAATT GTCAGCCCAA TAGTTTTTCA TTTTAAATTA 960

CNGAATTTTT TCATGTCTCT GGTTTTAGGA 990

(2) INFORMATION FOR SEQ ID NO: 133:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1720 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 133:

GTCTGATAAG CGACTGTGGT TATTCCCCTA AAGTTTACTT CAGCACTAAC ACTAGTGCTT 60

CCGCTGGAGT TTGCAGTTTT CCAGCTTTAT ACAGGATTTT CCTTTGACTG GAAGAGTCAA 120

GGATATAGAG ACTCAACAGT GACATTTATT GTACAACATC AAGGGGAATA GGATACTCAT 180

CAAACTGGGA TTATTCTTAT CAAAACATGG TCTTCTTTGA ATAAGAAAAA TACATAGTTG 240

GTTATTATGG ACTTAAAACT GTGTTAAATG GATATTCTGA TAAAATATTT GCTGCTCTGT 300

AGAGTGTGGA AAATCTGAGA ATATTAGCTT TACTCATCTT GAGCTTTGAG GATGTTCTCT 360

GTACGCCGAT GGTTTCATAT TAACTAAAAA AGCTGGGTAT TGTAAAATCT CATTTATAAA 420

AACTCAGATG AGAAGAAAAT TTTCTTTGAT GGTGAGACTG TTGTCTTAGT TCAGGAAATT 480

ATTTAATAAT CCTTTGTTAC CTGTGAATGA AGGAACTTTG TAATTCTGAT TTATCGTAAA 540

ACATGAGCCT TTCCAGAGTC AGCTTAGACA CTGTTGTCGC AAATAGCCAT GCTTTGCCTT 600

ATGCCAAGGA GGCCCAGAGG GAGGGCCTAG TCTTCCTCTG TTGCTGTACA TATATTGAAA 660

TGCTTTTTTT TTTTATTTTG CATTTGTTAT CTATAATGAG CTTTCTGAGC CCTGATATTA 720

TGTGAGACAA ACAGGAGTTA TTGATGTTAT ACACTCCCTT CCATTCAGGA TTTTCTGCTT 780

GGAGGGAAAT ATGTTGACCT TAGAGAATTG TGAATATTGT TGCAATTCTT GAATATATTA 840

CCATGTGAAT AATAGAGACT GTGTTGCTCT CTAGTATAAG CTATATTTAT TTTTGATTCA 900

TTTGAATTAC TAGTTATAAC TGGAGAAATT TTGTTACCTC TATCCTGGCT TGCCTGACTG 960

GCTGTATAAT AGCAGCAGCC TCTTTTAGAG CATCTTAATG AAAACATGGA TGAAAGGAAT 1020

TAATGATGAT ATCTGCAGAC TGCGTAGAAA ATGGCTTTTG TTCCCAGCGT TAACATTTTC 1080

TTCTCAATCA CATTTCAATG TTTGTGGAGA GTGGCAGATT CACACCAGAA ACACTAGGTG 1140

TTCATATCCA TAGCATGGAT GCAGAATAAG CAGTTGGGAG AGAAGCTTCT TCCTACCTGG 1200

TACTCCTCCC ATTCACCTCA GCCCAGCCCC AGACAGGCGT TAGCATTCAG TGTGGGCCCT 1260

CAGGCAGCCC TGAAGCCTGG CTGGGTCATC AGATGGGGGC AGCCTGTGAC GGGCACCAGC 1320

GGCCTGATTC CAGGGAAGAG TTCCTGGAGG GTGTTGGCTG TTTTTGTTAG CTCAGTTTTT 1380

TTCTGGGCTC CACCATTCCT AACTCCAGGT AGACAAGATA GATGTCACAC ACAACAATTT 1440

TAAAGTATTT TGCTTAGTGC ATTTTGTTTA TGATTGCAGT GTTTGTTTCT TATTTAATAG 1500

GCTTTTTACT TCATTCTATT AAATTTTAGT GTTTAGAAGA GGCGGGTACT GTCACTGTGT 1560

AAAATATGTA ATATTTTATA TGTTATACCA TGTCATATAT ACTTGCAATA TCAGACCTTG 1620

CATTCAATAT ACAATGCAAT TGACTCTTTG CAGACCTGCA TTTTTCAGTG AACAATAAAA 1680

AGATTGTCTG GCACTCCAAA AAAAAAAAAA AAAAAAAAAA 1720

(2) INFORMATION FOR SEQ ID NO: 134:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 705 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 134:

GGCACGAGGC CATCTGGGCT CATTCAGCAG GAAATAATGG AAAAAGCTGC AATATCCAGG 60

TGTTTACTAC AATCTGGAGG CAAGATCTTT CCTCAGTATG TGCTGATGTT TGGGTTGCTT 120

GTGGAATCAC AGACACTCCT AGAGGAGAAT GCTGTTCAAG GAACAGAACG TACTCTTGGA 180

TTAAATATAG CACCTTTTAT TAACCAGTTT CAGGTACCTA TACGTGTATT TTTGGACCTA 240

TCCTCATTGC CCTGTATACC TTTAAGCAAG CCAGTGGAAC TCTTAAGACT AGATTTAATG 300

ACTCCGTATT TGAACACCTC TAACAGAGAA GTAAAGGTAT ACGTTTGTNA AATCTGGGAA 360

GACTTGACTG CTATTCCATT TTGGGTATCA TATGTACCTT GATGAAGANG ATTAGGTTGG 420

GATACTTCAA GTGAAGCCTC CCACTGGAAA CAAGCTGCAG TTGTTTTAGA TAATCCCATC 480

CAGGTTGAAA TGGGAGAGGA ACTTGTACTC AGCATTCAGC ATCACAAAAG CAATGTCAGC 540

ATCACAGTAA AGCAATGAAG AGCAGTTTTC CAATGAAAAC TGTGTAAATA GAGCATCAAC 600

AAGTACAAAA TTCTTGTCTT AATTAGTGGG GGTATATAAA AATTCCTTGT AATGGTCAAA 660

TATTTTTTAA AATTGACATT AATAAAGCAT ATTTTAAAAG TTTCT 705

(2) INFORMATION FOR SEQ ID NO: 135:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 323 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 135:

AGCACACACC TCCTTTAGTT GCTCCTAAGG TCATGTTCAA CATTCGTGGA GTGCATTTTC 60

TGCTCAGGGA GCTTTCCCAG ACCCGGAATG TTTGGTGCTC ACAGACYCTG GCAAGGATCG 120

GTATTGCTGT TCCTCAGTTT TGCCTGGGGA AATGGAGGST CAGTGACGTT CAGTGACGTG 180

CCCAGAGTCA TGCCATTGGC GGGTGGCCCA GKGMTCCAGG TCTCCAGCAC CCCTCGGCCC 240

CCTCCTCACC AGGTCACATC ATCTCCTGGA TTAGAATCTG CTCACATAGT CTGTCCTGAA 300

AGGAAAAAAA AAAAAAAAAA AAC 323

(2) INFORMATION FOR SEQ ID NO: 136:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 582 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 136:

GGACGGAATG GTGCAACCCT CCTWAMTTTT CTKGKGCTGT TGACAACAGA GGGAGGGAGG 60

GAAAACATTT TTYGTGGGAG AATCCTACYT CTGCAGSGGA GCCCTTAAGC GATKGATTTT 120

GAATCTKGAC CCTTTACCAA CTAATTTTGA AGGAAGATAC CTTGGAAATA TTTGGCATTC 180

AGTGGGTTAC TGAAACAGCA TTAGTGAATT CATCTAGAGA ACTCTTTCAT TTATTCAGGC 240

AACAACTGTA CAACTTGGAA ACCTTGTTAC AGTCCAGTTG TGATTTTGGG AARGTATCAA 300

CTCTACACTG CAAAGCAGAC AATATTAGGC AGCAGTGTGT ACTATTTCTC CATTATGTTA 360

AAGTTTTCAT CTTCAGGTAT CTGAAAGTAC AGAATGCTGA GAGTCATGTT CCTGTCCATC 420

CTTATGAGGC TTTGGAGGCT CAGCTTCCCT CAGTGTTGAT TGATGAGCTT CATGGATTAC 480

TCTTGTATAT TGGACACCTA TCTGAACTTC CCAGTGTTAA TATAGGAGCA TTTGTAAATC 540

AAAACCAGAT TAAGGTTTGA CTGGTTTCAT TTGATTTTTA AG 582

(2) INFORMATION FOR SEQ ID NO: 137:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1021 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 137:

TTCGGCAGAG CCCTTGCGCG CTCTTGAATA CCTGCKTTCT GTAGCGCTAG TTCTCTTCAA 60

GATTTGCTTA GTGTCATTTC ATTTCGGTTT CTTTTCTCGC CATGTTTTTC TGTCGGAATT 120

ACGGTTCGTT TTGGTTCTAT GTACTCTCTA AAATGTTATC GTTTTTCATT TGTCTACTAA 180

TTTTCGTGCA TTTGTTACTA CTGAGTTTCT TAATATCTGA CTGGCCTCCG CCCACGGGCT 240

CTGCAGANCA TAAAATACTC AGGCTGATGG TAGTGCAGAG ACTCTCCCTC CTTGATCAGC 300

GCAAACGTTG GTCTGAGGCT TGAGGGATGG AGCAACATTT TCTTGGCTGT GTGAAGCGGG 360

CTTGGGATTC CGCAGAGGTG GCGCCAGAGC CCCAGCCTCC ACCTATTGTG AGTTCAGAAG 420

ATCGTGGGCC GTGGCCTCTT CCTTTGTATC CAGTACTAGG AGAGTACTCA CTGGACAGCT 480

GTGATTTGGG ACTGCTTTCC AGCCCTTGCT GGCGGCTGCC CGGAGTCTAC TGGCAAAACG 540

GACTCTCTCC TGGAGTCCAG AGCACCTTGG AACCAAGTAC AGCGAAGCCC ACTGAGTTCA 600

GTTGGCCGGG GACACAGAAG CAGCAAGARG CACCCGTAGA AKARGTGGGG CAGGCAGARG 660

AACCCGACAG ACTCAGGCTC CRGCAGCTTC CCTGGAGCAG TCCTCTCCAT CCYTGGGACA 720

GACAGCAGGA CACCGAGGTC TGTGACAGCG GGTGCCTTTT GGAACGCCGC CATCCTCCTG 780

CCCTCCAGCC GTGGCGCCAC CTCCCGGGTT TCTCAGACTG CCTGGAGTGG ATTCTTCGCG 840

TTGGTTTTGC CGCGTTCTCT GTACTCTGGG CGTGCTGTTC ACGGATCTGT GGAGCTAAGC 900

AGCCTTAGAT AGCAGCAGAA GGCTTTTTGG ATTCTCCTCC TTGAAAAGAT TCTCAGTTAC 960

CAAACGTCTC CACCTAGAAA ATAAAAATAC ATTAAGATGT TGANAAAAAA AAANAAAAAA 1020

A 1021

(2) INFORMATION FOR SEQ ID NO: 138:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1777 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 138:

CGGAAGATGA TGGCTTCAAC AGATCCATTC ATGAAGTGAT ACTAAAAAAT ATTACTTGGT 60

ATTCAGAACG AGTTTTAACT GAAATCTCCT TGGGGAGTCT CCTGATCCTG GTGGTAATAA 120

GAACCATTCA ATACAACATG ACTAGGACAC GAGACAAGTA CCTTCACACA AATTGTTTGG 180

CAGCTTTAGC AAATATGTCG GCACAGTTTC GTTCTCTCCA TCAGTATGCT GCCCAGAGGA 240

TCATCAGTTT ATTTTCTTTG CTGTCTAAAA AACACAACAA AGTTCTGGAA CAAGCCACAC 300

AGTCCTTGAG AGGTTCGCTG AGTTCTAATG ATGTTCCTCT ACCAGATTAT GCACAAGACC 360

TAAATGTCAT TGAAGAAGTG ATTCGAATGA TGTTAGAGAT CATCAACTCC TGCCTGACAA 420

ATTCCCTTCA CCACAACCCA AACTTGGTAT ACGCCCTGCT TTACAAACGC GATCTCTTTG 480

AACAATTTCG AACTCATCCT TCATTTCAGG ATATAATGCA AAATATTGAT CTGGTGATCT 540

CCTTCTTTAG CTCAAGGTTG CTGCAAGCTG GGAGCTGAGC TGTCAGTGGA ACGGGTCCTG 600

GAAATCATTA AGCAAGGCGT CGTTGCGCTG CCCAAAGACA GACTGAAGAA ATTTCCAGAA 660

TTGAAATTCA AATATGTGGA AGAGGAGCAG CCCGAGGAGT TTTTTATCCC CTATGTCTGG 720

TCTCTTGTCT ACAACTCAGC AGTCGGCCTG TACTGGAATC CACAGGACAT CCAGCTGTTC 780

ACCATGGATT CCGACTGAGG GCAGGATGCT CTCCCACCCG GACCCCTCCA GCCAAGCAGC 840

CCTTCAAGTT CTTTTATTTC TGGGTAACAG AAGTAGACAG ACAGGTTACT TGGTGTATCT 900

TCTGTTAAAG AGGATTGCAC GAGTGTGTTT TCCTCACACA CTTTGATTTG GAGAATTGGT 960

GCTAGTTGGC AATAGATAAC TCAGCGTAGA TAGTATTGCA AAAAGGGGAG GAAATACACA 1020

ACAATAATAA ATGTAAAAAC CTGCTATTCA ACATGCAGTT TTATTTCGAR GCCAAAAATC 1080

TAGAGCTTTC CCAAGATCCT GTTGCCTTAG GCACATNCAC ACTTCAACAG TGCACACTAT 1140

CCAACAGTGC ACACTATTCA ACAGTGCACA CTATTCAAAA GCGTAGACTA TTTTTTTGCA 1200

TGTTCAAGAT ATTTGTTTTG GTCTTATGTG TGTGTGAGAG AGAGAGATTC CTTTGACATT 1260

AAGGAGCATC AATGAGAAAA GATGATGAGG CAGGAATTAA TAAAGAAATG AAGTCGTGTG 1320

TGTTTGGTTG CCTGTCAGAG GGCACACAAT TTCATAAACA CCATGCCTGG ACAATTTGAT 1380

ATTAATATTT AACACCTCTG CATCTTTTTC TTAAAAAAGA ATATGGGCCA GATACAGTGG 1440

CTCACATTTG TAATCCCAGC ACTTTGGGGA GCCAAGTTAG CAGAATCCCT TGAGCACAGG 1500

AATCTGAAAC CAGCTTGGGC AACATAGTGA GATCCCATCT NTACAAAAAA CTTAAAAATT 1560

AGCCAGGCAT GATGGCACAT TCCTGTAGTC CTAGCTACTC AGGAGGCTAA GGTAGGAGGA 1620

TTGCCTGAGC CCAGGAGTTC AAGGCTGCAG TGAGCTAAGN ACGTGCCAGT ACACTCCAGC 1680

CTGAGCCACA AAGTGAGACC CTGTCTCGCA AAAAAAAAAN TTAAAAAGTC GGGGGGGGGC 1740

CCGGTACCCA AATCGCCGGA TATGATCGTA AACAATC 1777

(2) INFORMATION FOR SEQ ID NO: 139:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 643 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 139:

TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTGGG AATGAGAAAA TAACTTTATT 60

TTCATTGTGG GGAGCGGGCC GATGTCCAGC CTCAGAACTT CTGGAACTGC TTCTTGGTGC 120

CGGCAGCCTT GGTGACCTTG AGCACGTTGA AGCGCACTGT CTTGCTCAGA GGCCGGCACT 180

CGCCCACTGT GACGATGTCA CCGATCTGGA CGTCCCTGAA GCAGGGGGAC AGGTGTACAG 240

ACATGTTCTT GTGGCGCTTC TCGAAGCGGT TGTACTTGCG GATGTAGTGC AGATAGTCTC 300

GGCGGATGAC AATGGTCCTC TGCATCTTCA TCTTGGGTCA CCACGCCAGA GAGGATCCGC 360

CCTCGAATGG ACACATTACC AGTGAAGGGG CATTTCTTGT CAATGTAGGT GCCCCTCAAT 420

AGCCTCCTTG GGGTGTCTTT GAAGCCCAGA CCGATGTTCT TGTTAGTAAC CCGCGGGAGC 480

TTCTCCTTGC CAGTTTCTCC CAGCAGGACC CTCTTCTTGT TTTGAAAGAT GGTCGGCTGC 540

TTTTGGTAGG CACGCTCAGT CTGAATGTCC GCCATCTTCT CGTGCCGMAY TCCTGCAGCC 600

CGGGGGATCC ACTAGTTCTA GAGCGGCCGC ACCGCGGTGG AGC 643

(2) INFORMATION FOR SEQ ID NO: 140:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1220 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 140:

GGCACGAGGA TGATAGACCT ACTGGAGGAA TACATGGTTT ACAGGAAGCA TACCTACATR 60

AGGCTTGATG GCTCATCCAA GATCTCGGAG AGGCGAGACA TGGTTGCTGA TTTTCAGAAC 120

AGGAATGACA TCTTTGTGTT CCTGTTAAGC ACACGAGCTG GAGGACTGGG TATCAATCTC 180

ACTGCTGMAG ACACAGTGCA TTTTCTATGA TAGCGACTGG AACCCCACTG TGGACCAGCA 240

GGCCATGGAC AGGGCCCACC GCTTAGGGCA GACAAAGCAG GTTACTGTGT ACCGGCTCAT 300

CTGTAAAGGC ACCATTGAAG AACGCATTCT GCAAAGAGCC AAGGAGAAGA GTGAGATTCA 360

GCGGATGGTG ATTTCAGGTG GGAACTTCAA ACCAGATACC TTGAAACCCA AAGAGGTGGT 420

TAGTCTTCTT CTAGACGACG AAGAGTTGGA GAAGAAACGT ATGTACTCTA AACCTCTATA 480

CACTCCCCTC ACGTATCTGA GAATGGAAGA GGTACTTGGS TGTGTGCCAA GGGTTAGGCA 540

AAGCCAGAGG CTGTATTTAG GGAAAGTATT TTTGTGCTCA TATTTTATAT AAAAACCCAA 600

ACAAGAATGT GTTTGTAGGC CAGGCGTGGT GGCTCGCGCC TCTAGTCTCA GCATTTCGGG 660

ARGCCAAAGT GGGCAGATCA CCTGARGTCA GGARTTTGAG TTTGARACCA GCCTGGCCMA 720

CGTTGTGAAA CCCCACCTCT ACTARGARTA CSGAAAATTG GTTGGGCATG GTGGCGGGCA 780

CCTGTAATTC CAGCACTTTG GGAGGCTGGG GCAGAANAAT TGCTTGAGCC CAGGAGGTGG 840

AGATTGCGGT GAGCCGAGAT YGTGCCATTG CAMTCCAGCC SGGGCAATAA GAGTGAAAYT 900

CCATCTTTTA AAAACAAACA AAAACAAAAA ACACAAGACG GCTCACACCT GTAATCCCAG 960

CACTTTGGGA RGCCGARGCA GGTGGATCAC GARGTCAGGA GTTCCAAGAC TAGCCTGGCC 1020

AACCTGGTGA AGCCCCGTCT CTACTAAAAA TACMAATATT AGTCGGGCGT GGTGGTGGGC 1080

ACGTGTAATC CCAGCTACTC GGGAGGCTGA GGCAGGAGAA TCCCTTGAAG CTAGGAGGCA 1140

GAGGTTGCAG TGAGCCAGGA TCGTGCCATT GCACTCCAGC CTGGACAACA AGAGCAAGAT 1200

TCCATCTCAA AAAAAAAAAA 1220

(2) INFORMATION FOR SEQ ID NO: 141:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 721 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 141:

AATTCGGCAC GAGCCAGGTT AGCCGGAAGG GCAGCTCTCC AGGCCCTGCC CACCCCACAG 60

GGGGCTCCTT ATGCACAGCG GGGCGTCTCC TTGTGGCCAT AGAAACGGAA CTGGCTCTTT 120

TCAACAGTGC TGCAAGAGGA TGGTTATTTA ACGCTGGCCC CCAAGGAGGA AAGGCACAGA 180

CYTTCCTCCC TCCTGGAACA TCCAAGGGCA CTGGATCCTC TGTGTCCCTC TGAGATGGGG 240

TGCCACTCCA GCAAGAGCAC CACGGTGGCA GCTGAGTCCC AGAAGCTTGA AGAAGAGYGC 300

GAGGGAAGAG AGCCAGGTCT GGAGACCGGC ACCCAGGCAG CAGACTGCAA GGATGCCCCG 360

CTGAAGGATG GAACCCCTGA GCCAAAGAGC TGAAATGCCT CTCTCCAGAG TCGGACCCTC 420

ACCTCYTTCC TGGAACTGCC TTTGGCCCCA GAACCATGAG ACAATCCCCA CCCTGAGAAG 480

CTCCGATCAC TGGGAGGAGA GAGAAAGCCT CCAGCTTTGG GATTCAGGCT TCAGAAGTTT 540

TTAGCAGCCT TTGCTCATTG GAGAGGTGGG GAAAGGATAA AGTTCTTATA AGGAAATCCC 600

TAATTTCCCC CAGCTCCTCC CCNCCNGAAG AAGGAACNAA AGAAAGTTCC TTCCACACGT 660

TTTGTTGGAA ACTTTTCCCT TGCCAACTTT CCTTGGATTG CCAGAACAAA GCCCTCCAGA 720

A 721

(2) INFORMATION FOR SEQ ID NO: 142:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1468 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 142:

ATGAATTAAT GTTTATAAAT GACTGTACTG AATTTAAAAC CGTACAGTTT CATTTGCATT 60

TTGACATTAC TTTATTATAC ATTTTGCATT TAAAAGGCTG CACCAGTTGG CTTTTCTTCT 120

GTTTTATTCT CAAAATATAG AGATTCTGTG ATTTATTTGC CCTGTTTATG GATTAAAAAG 180

AAAATTCTAA TATAAAGCAT TTCAATAGGA TGCATAGGTA TATTACGTTT TTTAAATGCT 240

TTAGATCTGT GATTCTTGAC TTACTATTTA TTTTATCCCC TTTAAGTCAG GGATGCTTTA 300

TTCTATTTTA AAGCACTTAT GAGTTACATG TTGTAATCAA GTTTGCACAA TATATTTATC 360

TATATGAGGA ACCCATAAAT GAATAGCTAA TTTTTAAAAT GCCATTAAAA TGCATGAAAT 420

KCTTATTAAA ACCTTACTAT ACTATTTCTT CAAGGCAAGT AAATTGACCA TGRGRAAAGR 480

ACACAGTTAT TAAACACTGT TGACAGGAAA ATTCTCCTTG ATAACATAGG ACAATTAATG 540

GAAAAAAAAA TTCTCATTAT TTGCAAAGAA TGAACAAGTT AATGAACAAA CAAACTAGAT 600

TTGGTATGTT TTCAGCTTTT GTATCATGTT TAATTGTTTA ATTTGGTTGA AAAACTGCAG 660

TTGAGAAATC AGATAGCAAT ATAGACATTC ACAGCAGCTC TGTGGATACC ATGTAATTGT 720

CAGGTAATTT CAGAATGTTG AAAATTATTC AGTGCAGCCC TCATAGTATC ATACTTGAAG 780

AAATTGATTA CAGTTCCACT AAATTGTTGA AGATAAATTA TTTTTAAAGG TTATGAAAAC 840

TAAGTTATAT TAATTCATAT GTTTGATTTT TAAATCCCAC CTCCTCAAGC TATCCAATTT 900

NCTGACTTTG AAAATAACCA TGAGAGATGC CACATTTCTC TCTGGGAAAC TACCACTCAA 960

AGAATAATTG TTAAAAATTA AGCTTTTAGG TATTAGAAGC TGTTATAAAG TATAAAATTA 1020

AGATATAAGC AGATCACATG TAAATCATTC CTAAAGCACA AGAAAAGAAT GTGCCTTGAT 1080

GTACATATAT TACTAAGTTG CCTCTCCCAG TTTACTTTAA AAATGGCTTT AAGGATAAAG 1140

AATAAATGTG ATAGCTGTGC ATGCATTATA TATTTGCATT TGCAAATTTC CCATTGTTTT 1200

AACAGCTGTG TGGCTGACTT TCAATTTTAA GACGTGAATT GACATACAGC CCATAACTTT 1260

ATAATGGCTG CTCATTTATC TTATCTTTCA GTTAGTGGAA AAACATTTCA ACCTGACTAA 1320

AATTTGGAAT TGTGTCTTTT ATGTTCCATC CTCTGTTGTT ACTAGATTTA GTTTAAAAAT 1380

TGTGTATGAC CATTAATGTA TGTCATAAAC ATGTAAATAA AAGATGTTGA ATCTTGTTGA 1440

AAAGCAWRAA AAAAAAAAAA AAACTCGA 1468

(2) INFORMATION FOR SEQ ID NO: 143:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 300 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 143:

TGAATTTTTT GCCAAACTTA GTAACTCTGT TAAATATTTG GAGGATTTAA AGAACATCCC 60

AGTTTGAATT CATTTCAAAC TTTTTAAATT TTTTTGTACT ATGTTTGGTT TTATTTTCCT 120

TCTGTTAATC TTTTGTATTC RCTTATGCTC TCGTACATTG AGTACTTTTA TTCCAAAACT 180

AGTGGGTTTT CTCTACTGGA AATTTTCAAT AAACCTGTCA TTATTGCTTA CTTTGATTAA 240

AAAAAAAAAA AAAAAAAAAA AAACCCCNAG GGGGGGGCCG GGTNCCCAAT CCCCCCCAAA 300

(2) INFORMATION FOR SEQ ID NO: 144:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2243 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 144:

TGCCTCCCTT CCTGCAGATT GTGGACAGTA GTTCCTCAGC CTGCACCCTG GATTCCTTCT 60

TCCCCTTCCT AGCTCCATGG GACTCGCCCC AAGACTGTGG CTTCAAGGAC CACCAGCCCC 120

TTACTCTTCA AGCCCTGACT GTGGAGTTGG TAGATGCCTC TGATCCTCAG TATTCTCTCT 180

GGCAATGTTC CACGGCTTCT CCTTCCTGGG AGCTGGCTCC ATAACTTGAT TTTCCCCAAA 240

CGTGTTGCAA TCCCTGCTGC CCCTTAGCCA CCCAGGGTCT TGTGTGGGTA TGAGTGTAGA 300

GGATGGGGGT ATGCCAGGCC TGGGCCGTCC CAGGCAGGCC CGCTGGACCC TGATGCTACT 360

CCTATCCACT GCCATGTACG GTGCCCATGC CCCATTGCTG GCACTGTGCC ATGTGGACGG 420

CCGAGTGCCC TTYCGGCCCT CCTCAGCCGT GCTGCTGACT GAGCTGACCA AGCTACTGTT 480

ATGCGCCTTC TCCCTTCTGG TAGGCTGGCA AGCATGGCCC CAGGGGCCCC CACCCTGGCG 540

CCAGGCTGCT CCCTTCGCAC TATCAGCCCT GCTCTATGGC GCTAACAACA ACCTGGTGAT 600

CTATCTTCAG CGTTACATGG ACCCCAGCAC CTACCAGGTG CTGAGTAATC TCAAGATTGG 660

AAGCACAGCT GTGCTCTACT GCCTCTGCCT CCGGCACCGC CTCTCTGTGC GTCAGGGGTT 720

AGCGCTGCTG CTGCTGATGG CTGCGGGAGC CTGCTATGCA GCAGGGGGCC TTCAAGTTCC 780

CGGGAACACC CTTCCCAGTC CCCCTCCAGC AGCTGCTGCC AGCCCCATGC CCCTGCATAT 840

CACTCCGCTA GGCCTGCTGC TCCTCATTCT GTACTGCCTC ATCTCAGGCT TGTCGTCAGT 900

GTACACAGAG CTGCTCATGA AGCGACAGNG GCTGCCCCTG GCACTTCAGA ACCTCTTCCT 960

CTACACTTTT GGTGTGCTTC TGAATCTAGG TCTGCATGCT GGCGGCGGCT CTGGCCCAGG 1020

SCTCCTGGAA GGTTTCTCAG GATGGGCAGC ACTCGTGGTG CTGAGCCAGG CACTAAATGG 1080

ACTGCTCATG TCTGCTGTCA TGAAGCATGG CAGCAGCATC ACACGCCTCT TTGTGGTGTC 1140

CTGCTCGCTG GTGGTCAACG CCGTGCTCTC AGCAGTCCTG CTACGGCTGC AGCTCACAGC 1200

CGCCTTCTTC CTGGCCACAT TGCTCATTGG CCTGGCCATG CGCCTGTACT ATGGCAGCCG 1260

CTAGTCCCTG ACAACTTCCA CCCTGATTCC GGACCCTGTA GATTGGGCGC CACCACCAGA 1320

TCCCCCTCCC AGGCCTTCCT CCCTCTCCCA TCAGCAGCCC TGTAACAAGT GCCTTGTGAG 1380

AAAAGCTGGA GAAGTGAGGG CAGCCAGGTT ATTCTCTGGA GGTTGGTGGA TGAAGGGGTA 1440

CCCCTAGGAG ATGTGAAGTG TGGGTTTGGT TAAGGAAATG CTTACCATCC CCCACCCCCA 1500

ACCAAGTTCT TCCAGACTAA AGAATTAAGG TAACATCAAT ACCTAGGCCT GAGAAATAAC 1560

CCCATCCTTG TTGGGCAGCT CCCTGCTTTG TCCTGCATGA ACAGAGTTGA TGAAAGTGGG 1620

GTGTGGGCAA CAAGTGGCTT TCCTTGCCTA CTTTAGTCAC CCAGCAGAGC CACTGGAGCT 1680

GGCTAGTCCA GCCCAGCCAT GGTGCATGAC TCTTCCATAA GGGATCCTCA CCCTTCCACT 1740

TTCATGCAAG AAGGCCCAGT TGCCACAGAT TATACAACCA TTACCCAAAC CACTCTGACA 1800

GTCTCCTCCA GTTCCAGCAA TGCCTAGAGA CATGCTCCCT GCCCTCTCCA CAGTGCTGCT 1860

CCCCACACCT AGCCTTTGTT CTGGAAACCC CAGAGAGGGC TGGGCTTGAC TCATCTCAGG 1920

GAATGTAGCC CCTGGGCCCT GGCTTAAGCC GACACTCCTG ACCTCTCTGT TCACCCTGAG 1980

GGCTGTCTTG AAGCCCGCTA CCCACTCTGA GGCTCCTAGG AGGTACCATG CTTCCCACTC 2040

TGGGGCCTGC CCCTGCCTAG CAGTCTCCCA GCTCCCAACA GCCTGGGGAA GCTCTGCACA 2100

GAGTGACCTG AGACCAGGTA CAGGAAACCT GTAGCTCAAT CAGTGTCTCT WTAACTGCAT 2160

AAGCAATAAG ATCTTAATAA AGTCTTCTAG GCTGTAGGGT GGTTCCTACA ACCACAGCCA 2220

AAAAAAAAAA AAAAAAACTC GAG 2243

(2) INFORMATION FOR SEQ ID NO: 145:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1082 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 145:

GCCAAGCTCT AATACGACTC ACTATAGGGA AAGCTGGTAC GCCTGCAGKT ACCGGTTCCG 60

GGAATTCCCG GGTCGACCCA CGCGTCCGCT TCCGTGTGTC AAAATCCTCA CCTCCTTCAT 120

AACCATCTCC CACAATTAAT TCTTGACTAT ATAAATTTAT GGTTTGATAA TATTATCAAT 180

TTGTAATCAA TTGAGATTTC TTTAGTGCTT GCTTTTCTGT GACTCAACTG CCCAGACACC 240

TCATTGTACT TGAAAACTGG AACANCTTGG GAATGCCATG GGGTTTGATA ATCTGCCAGG 300

GACATGAAGA GGCTCAGCTT CCTGGGACCA TGACTTTGGC TCAGCTGATC CTGNACATGG 360

GAGAACAACC ACATTTTTCT TTGTGTGTGC TTCTAGCAGC TGTTCGGGAG GACCKTGACC 420

CAAYAGTGTT CCCATGCTGT TTCTTGTGAA ATGCTCTCGG CTATGTAGCA GCTTTTGATT 480

CCCTGCATAC CCTAGGCTGC TGCCCCTATC CTGTCCCTTG TTTATAACAT TGAGAGGTTT 540

TCTAGGGCAC ATACTGAGTG AGAGCAGTGT TGAGAAGTCG GGGAAAATGG TGACTACTTT 600

TAGAGCAAGG CTGGGCATCA GCACCTGTCC AGCTCTACTT GTGTGATGTT TCAGGAACTC 660

AGCCCCTTTT TCTGCCTAGG ATAAGGAGCT GAAAGATTAA CTTGGATCTY CTAATGGTCC 720

AAATCTTTTG GTCACAATAA AGAGTCTCCA AATTAGAGAC TGCATGTTAG TTCTGGATGG 780

ATTTGGTGGC CTGACATGAT ACCCTGCCAG CTGTGAGGGG ACCCCGTTTT TAAGATGCAT 840

GGCCAAGCTC TCTGCAAATG GAAATGCTTA CACTGGGTGT TGGGGATGTT TGCTACCTCC 900

TGCTATTTTT GTGGTTTTGG TTCTCCCACT ATGGTAGGAC CCCTGGCCAG CATTGTGGCT 960

TGTCATGTCA GCCCCATTGA CTACCTTCTC ATGCTCTGAG GTACTACTGC CTCTGCAGCA 1020

CAAATTTCTA TTTCTGTCAA TAAAAGGAGA TGAAAATAAA AAANAAAAAA AAAAAACTCG 1080

NG 1082

(2) INFORMATION FOR SEQ ID NO: 146:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4313 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 146:

CAAGCTGGTT TGAAACTAGG GGTCGGGCTC GGCCGTCGTC GTTGTTTGTC GCCGCATCCC 60

CGCTTCCGGG TTAGGCCGTT CCTGCCCGCC CCCTCCTCTC CTCCCTTCGG ACCCATAGAT 120

CTCAGGCTCG GCTCCCCGCC CGCCGCAGCC CACTGTTGAC CCGGCCCGTA CTGCGGCCCC 180

GTGGCCACCA TGTCCCTGCA CGGCAAACGG AAGGAGATCT ACAAGTATGA AGCGCCCTGG 240

ACAGTCTACG CGATGAACTG GAGTGTGCGG CCCGATAAGC GCTTTCGCTT GGCGCTGGGC 300

AGCTTCGTGG AGGAGTACAA CAACAAGGTT CAGCTTGTTG GTTTAGATGA GGAGAGTTCA 360

GAGTTTATTT GCAGAAACAC CTTTGACCAC CCATACCCCA CCACAAAGCT CATGTGGATC 420

CCTGACACAA AAGGCGTCTA TCCAGACCTA CTGGCAACAA GCGGTGACTA TCTCCGTGTG 480

TGGAGGGTTG GTGAAACAGA GACCAGGCTG GAGTGTTTGC TAAACAATAA TAAGAACTCT 540

GATTTCTGTG CTCCCCTGAC CTCCTTTGAC TGGAATGAGG TGGATCCTTA TCTTTTAGGT 600

ACCTCAAGCA TTGATACGAC ATGCACCATC TGGGGGCTGG AGACAGGGCA GGTGTTAGGG 660

CGAGTGAATC TCGTGTCTGG CCACGTGAAG ACCCAGCTGA TCGCCCATGA CAAAGAGGTC 720

TATGATATTG CATTTAGCCG GGCCGGGGGT GGCAGGGACA TGTTTGCCTC TGTGGGTGCT 780

GATGGCTCGG TGCGGATGTT TGACCTCCGC CATCTAGAAC ACAGCACCAT CATTTACGAA 840

GACCCACAGC ATCACCCACT GCTTCGCCTC TGCTGGAACA AGCAGGACCC TAACTACCTG 900

GCCACCATGG CCATGGATGG AATGGAGGTG GTGATTCTAG ATGTCCGGGT TCCTGCACAC 960

CTGTSGCCAG GTTAAACAAC CATCGAGCAT GTGTCAATGG CATTGCTTGG GCCCCACATT 1020

CATCCTGCCA CATCTGCACT GCAGCGGATG ACCACCAGGC TCTCATCTGG GACATCCAGC 1080

AAATGCCCCG AGCCATTGAG GACCCTATCC TGGCCTACAC AGCTGNAAGG WGAGATCAAC 1140

AATGTGCAGT GGGCATCAAC TCAGCCCGAA YTGTCGCCAT CTGCTACAAC AACTGCCTGG 1200

AGATACTCAG AGTGTAGTGT TGGTGGCGCT GTGCCCACGA GGCAGGGGCT TTTGTATTTC 1260

CTGCCTCTGC CCCACCCCCA AAGTAAGAAG AAACATGTTT CCAGTGGCCA GTATGTCTTT 1320

CATTGCTTTG CACCCACTGT TACCAGAAGC TGCTCTAGGA GTTCCTGGCC AGTCACCCCA 1380

TCGCCCTCTG TGGCAGACTC AGTGCTGTGT GGCGCCTCCT CAGCCCAGGG CTGAGTTTTA 1440

AGATTTTCTC TCCTTTCCTC TTCTCCTTTG GTTCCTCAAT TAAAAAATGT GTGTATATTT 1500

GTTTGTCAGG CGTTGTGTTG AGGAGCAGTT CACGCACTGG CTGTGTCTAT TCCTCTGCCC 1560

AGGTGTCTCT GTTTGCTGCC CAAKGYWKKT TTTCATGTCT CGTCCATGTC CATGTTCGTG 1620

TTAGCACTWA CGTGGGAACA AATACCAATT TGTCTTTTCT CCTAGTATCA GTGTGTTTAA 1680

CAAATTTTAA CTTTGTATAT TTGTTATCTA TCAGGCTAAT TTTTTTATGA AAAGAATTTT 1740

ACTCTCCTGC TTCATTTCTT TGTCTTATAG TCCTCCCTCT TTGCACCTTC TTCTCTTCCC 1800

TCAGTGCCTG GAGCTGGTAC TGGGCCCCTG GCCCCATGAG CAGTTTGCCT TCTTGAGTCA 1860

CTGCCTGTGT AGTACATACC TGACCGGGAG TCCAAACCAC CTTGGTGCTC TGAAGTCCAC 1920

TGACTCATCA CACCTTTCTT AGCCTGGCTC CTCTCAAGGG CATTCTGGGC TTGTAAACAG 1980

ACATAGGAAG CCTCTGTTTA CCCTGAAGCA CCACTGTCCA GCCCATTGGT TCCCACTGGC 2040

AGCATGGTAG AGCTGAGAGA AACAGGCTCT CAGGGTACCT GACTTGAGGG GAATCGTTTC 2100

ATGAAGCTGA ACTTCAAGCA TATTTCCAGT ACATTCTTTC AGAGTCTGTT TTTCCATCCA 2160

AATATAAGCC CCAGGCCATT CCACTTAGTG TCTTTTCAAT GATAGGCAAG AATGATATCT 2220

GAGTTGAACT TCGGTGCTTC TGTTGTTTGA GTTTACTGTG CCTGGTGGTA TATTGGGCAT 2280

TCTTTGGATT GAGTGTTCTG AGGTGAGAGA GTCTTCCCGA GGCATCCTGT CTGTGCTTCC 2340

AACCCTGAAC AAGACCTTAC ATGAGAGATG GACTGATGGA CTGCGGCAAT CCTGGGCTGT 2400

CAAGTGGATA GATAGTTAAA AAGCATTATA CTGTGGGTAA TGAAAAGGGA GGAAAAAAAA 2460

AGAAGGAAAA GGAATTATAG ACCCCCAGGG TCAGCCAGTT AAGAGCTCTA CCCACACCTG 2520

TCAACCCCTC TCTCCCCCAG TTTAGGTTCT GAGCAGTATT GGACTTGTAG CCTGCAGTTG 2580

TCTTTTGACT TGCAGGCCGC AGTGTCTTTC TGTTATGTGA ATGAGTTCCA TGGAGGGGCA 2640

TATGTGTGAT TCCACCGTTA GATGAGCCCT TGGGGCAGGC AGTTTGGGAT GTGCTCTTGG 2700

GGGAAAGTTG GCTGTTTCCT TGCGCTCTGC TCCTACCCGA AGTTTTTAAG TCCCTCTGAA 2760

TTGCTCATCT GAGATTAGTA GAGTAGCAGG CCTGAAGGAT GATGGTTTTG TCCTCTTTGG 2820

TTCTCACCTG CTTGAGAAGT AAAACAGTAA CTTTGTTCTT CTGGGCCCTT AAGCTTTTTT 2880

GGTTAAGTCT TCCTTTTCAG AAGTAGATGT CATTATATGC CAAAAGTCTA GCTCTTTGCT 2940

TTACCATACA GGGACCTGTC CCAAAGAAAA AGGCTCTTTT TTTAGCCAGC ATATTTCCCC 3000

TTCTACCCTT TTACTTTGTT GTTCTGATTT TAGGACTCTG GCTGGCCATG TGCTTGTGGT 3060

TGCCTCTCCT GCATTTGCCA CTGGATTTGC ACTGCATCGT TTGGAGATAC AAAGCGAGCA 3120

GTTCTTGGTC AGAACCCTCC TCTGCTTTTC ATTGTGTTTG ATAATGGTTA CTGGGTCCTT 3180

CTCTCAAGGG TAGCAAGGCC AAGCTGATGG CTGCTTGTTT AGGAGGCCAT CAGTTCCTTC 3240

CTGTGGAGAA GGGTCTGAAA TGGAAGTCAG TGGTAGAAGG GGCTGGTCTG CTGGGCAGGG 3300

CTTACATCCA CTGAGTTCTA AGATTCCTTT CCTGATCTGC ACCTACGCCT GGTCTGTATG 3360

GTGGAATTTG TCAGCTGGAA CTCAGAAACA ACAACTTGAA AAAAAAATAA TAATTAGAAC 3420

ATATTTGCAT AAGATAGCTA TTTACTCTGG AAACCAACAA CTTTTGAGAT TTCCCTTGCC 3480

CTGTGGACGC CCAGCTCCTG TCATCCTTCC TTAGGTCCTG CAGTACAGTC TTCCCCTGAA 3540

TGCCACCGGG GACCCAGGGG GACTCCACCC CCCTAAGCAA GCACACACAT ACTCACAGTT 3600

GATGAGTTGC TGGTCTTTGA GTCCCAGCTC TCTTACCCTC CCTTTACTCC ACCAGCCCGA 3660

CGACCCATGA CTGAGGAGGG GATTTCTACA GTCTCAGGAT TTAGAAAGTC TGTAAGCCAT 3720

CCATGCTCCA GAAAGCACCG ATCTGTTGTA GTTGCAAAAA CAACTCTGTA ATTTGTTGAG 3780

GTTCTCAAAC TGACAGCCAG CGAGACTGGG TGGGAGGCCC TGGATCTGTT CTCCCTGACT 3840

GCGGGAGGAG CAGCCACTAG GACTTTAGCA GGAAGCCCAC ATGGAGGCTC CGCCAGGCTG 3900

TGGCCCAGCT GGTGATGGCC CTTTTGCTCC TGGCAGCCTG AGGCACAGCT GCCTGTATTG 3960

TCCTCATCTG TTCTGACTGA AGGATGGAGG TGCTGAATAA ATTAGGCCTC AGGCNTCTAC 4020

CACCAGAGAG CTGGAGAATG GGTCCACGTC ATTCAAGGAC CTGAATTTTT TATGCTCAGG 4080

AGCATTGGAA TCCTCTTCTT CCAGGGAGGA ATTAGCCTGC AAGGTTAGGA CTTGAAGAGG 4140

GAAGGTATTT AATAACTGGG CGAGGATGGG TGTGGTGGCT CACACCTGTA ATCCCAGCAT 4200

TTTGGGAGGC TGAGGTGGCC AGATCCCAAG GTCAGAAGAT CGAGACCATC CTGGCTAACA 4260

TGGTGAAACC CCATCTCTAC TAAAAATACA AAATTAAATT GGCCGGGCGT GAA 4313

(2) INFORMATION FOR SEQ ID NO: 147:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1183 base pairs
(B) TYPE: nucleic acid